P · A · R · T · 2

PHYSICAL OPTICS

CHAPTER 2
INTERFERENCE

John E. Greivenkamp
Optical Sciences Center
University of Arizona
Tucson, Arizona

2.1 GLOSSARY

A	amplitude
$\mathbf{E}$	electric field vector
$\mathbf{r}$	position vector
x, y, z	rectangular coordinates
ϕ	phase

2.2 INTRODUCTION

Interference results from the superposition of two or more electromagnetic waves. From a classical optics perspective, interference is the mechanism by which light interacts with light. Other phenomena, such as refraction, scattering, and diffraction, describe how light interacts with its physical environment. Historically, interference was instrumental in establishing the wave nature of light. The earliest observations were of colored fringe patterns in thin films. Using the wavelength of light as a scale, interference continues to be of great practical importance in areas such as spectroscopy and metrology.

2.3 WAVES AND WAVEFRONTS

The *electric field vector* due to an electromagnetic field at a point in space is composed of an amplitude and a phase

$$\mathbf{E}(x, y, z, t) = \mathbf{A}(x, y, z, t)e^{i\phi(x,y,z,t)} \tag{1}$$

or

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{A}(\mathbf{r}, t)e^{i\phi(\mathbf{r},t)} \tag{2}$$

where $\mathbf{r}$ is the position vector and both the amplitude $\mathbf{A}$ and phase ϕ are functions of the spatial coordinate and time. As described in Chap. 5, ''Polarization,'' the polarization state of the field is contained in the temporal variations in the amplitude vector.

This expression can be simplified if a linearly polarized monochromatic wave is assumed:

$$\mathbf{E}(x, y, z, t) = \mathbf{A}(x, y, z)e^{i[\omega t - \phi(x,y,z)]} \tag{3}$$

where ω is the angular frequency in radians per second and is related to the frequency ν by

$$\omega = 2\pi\nu \tag{4}$$

Some typical values for the optical frequency are 5×10^{14} Hz for the visible, 10^{13} Hz for the infrared, and 10^{16} Hz for the ultraviolet. Note that in the expression for the electric field vector, the time dependence has been eliminated from the amplitude term to indicate a constant linear polarization. The phase term has been split into spatial and temporal terms. At all locations in space, the field varies harmonically at the frequency ω.

Plane Wave

The simplest example of an electromagnetic wave is the *plane wave.* The plane wave is produced by a monochromatic point source at infinity and is approximated by a collimated light source. The complex amplitude of a linearly polarized plane wave is

$$\mathbf{E}(x, y, z, t) = \mathbf{E}(\mathbf{r}, t) = \mathbf{A}e^{i[\omega t - \mathbf{k}\cdot\mathbf{r}]} \tag{5}$$

where $\mathbf{k}$ is the wave vector. The wave vector points in the direction of propagation, and its magnitude is the wave number $k = 2\pi/\lambda$, where λ is the wavelength. The wavelength is related to the temporal frequency by the speed of light v in the medium:

$$\lambda = \frac{v}{\nu} = 2\pi\frac{v}{\omega} = \frac{c}{n\nu} = 2\pi\frac{c}{n\omega} \tag{6}$$

where n is the index of refraction, and c is the speed of light in a vacuum. The amplitude $\mathbf{A}$ of a plane wave is a constant over all space, and the plane wave is clearly an idealization.

If the direction of propagation is parallel to the z axis, the expression for the complex amplitude of the plane wave simplifies to

$$\mathbf{E}(x, y, z, t) = \mathbf{A}e^{i[\omega t - kz]} \tag{7}$$

We see that the plane wave is periodic in both space and time. The spatial period equals the wavelength in the medium, and the temporal period equals $1/\nu$. Note that the wavelength changes with index of refraction, and the frequency is independent of the medium.

Spherical Wave

The second special case of an electromagnetic wave is the spherical wave which radiates from an isotropic point source. If the source is located at the origin, the complex amplitude is

$$E(r, t) = (A/r)e^{i[\omega t - kr]} \tag{8}$$

where $r = (x^2 + y^2 + z^2)^{1/2}$. The field is spherically symmetric and varies harmonically with time and the radial distance. The radial period is the wavelength in the medium. The amplitude of the field decreases as $1/r$ for energy conservation. At a large distance from the source, the spherical wave can be approximated by a plane wave. Note that the vector characteristics of the field (its polarization) are not considered here as it is not possible to describe a linear polarization pattern of constant amplitude that is consistent over the

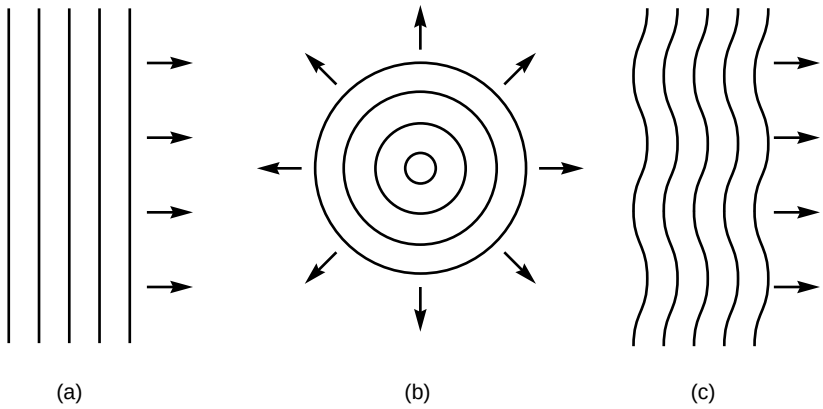


FIGURE 1 Examples of wavefronts: (*a*) plane wave; (*b*) spherical wave; and (*c*) aberrated plane wave.

entire surface of a sphere. In practice, we only need to consider an angular segment of a spherical wave, in which case this polarization concern disappears.

Wavefronts

Wavefronts represent surfaces of constant phase for the electromagnetic field. Since they are normally used to show the spatial variations of the field, they are drawn or computed at a fixed time. Wavefronts for plane and spherical waves are shown in Fig. 1*a* and *b*. The field is periodic, and a given value of phase will result in multiple surfaces. These surfaces are separated by the wavelength. A given wavefront also represents a surface of constant optical path length (OPL) from the source. The OPL is defined by the following path integral:

$$\text{OPL} = \int_S^P n(s)\, ds \tag{9}$$

where the integral goes from the source S to the observation point P, and $n(s)$ is the index of refraction along the path. Variations in the index or path can result in irregularities or aberrations in the wavefront. An aberrated plane wavefront is shown in Fig. 1*c*. Note that the wavefronts are still separated by the wavelength.

The local normal to the wavefront defines the propagation direction of the field. This fact provides the connection between wave optics and ray or geometrical optics. For a given wavefront, a set of rays can be defined using the local surface normals. In a similar manner, a set of rays can be used to construct the equivalent wavefront.

2.4 INTERFERENCE

The net complex amplitude is the sum of all of the component fields,

$$\mathbf{E}(x, y, z, t) = \sum_i \mathbf{E}_i(x, y, z, t) \tag{10}$$

and the resulting field intensity is the time average of the modulus squared of the total complex amplitude

$$I(x, y, z, t) = \langle |\mathbf{E}(x, y, z, t)|^2 \rangle \tag{11}$$

where $\langle\ \rangle$ indicates a time average over a period much longer than $1/\nu$. If we restrict ourselves to two interfering waves $\mathbf{E}_1$ and $\mathbf{E}_2$, this result simplifies to

$$I(x, y, z, t) = \langle|\mathbf{E}_1|^2\rangle + \langle|\mathbf{E}_2|^2\rangle + \langle\mathbf{E}_1 \cdot \mathbf{E}_2^*\rangle + \langle\mathbf{E}_1^* \cdot \mathbf{E}_2\rangle \tag{12}$$

or

$$I(x, y, z, t) = I_1 + I_2 + \langle\mathbf{E}_1 \cdot \mathbf{E}_2^*\rangle + \langle\mathbf{E}_1^* \cdot \mathbf{E}_2\rangle \tag{13}$$

where I_1 and I_2 are the intensities due to the two beams individually, and the (x, y, z, t) dependence is now implied for the various terms.

This general result can be greatly simplified if we assume linearly polarized monochromatic waves of the form in Eq. (3):

$$\mathbf{E}_i(x, y, z, t) = \mathbf{A}_i(x, y, z)e^{i[\omega_i t - \phi_i(x,y,z)]} \tag{14}$$

The resulting field intensity is

$$I(x, y, z, t) = I_1 + I_2 + 2(\mathbf{A}_1 \cdot \mathbf{A}_2)\cos[(\omega_1 - \omega_2)t - (\phi_1(x, y, z) - \phi_2(x, y, z))] \tag{15}$$

The interference effects are contained in the third term, and we can draw two important conclusions from this result. First, if the two interfering waves are orthogonally polarized, there will be no visible interference effects, as the dot product will produce a zero coefficient. Second, if the frequencies of the two waves are different, the interference effects will be modulated at a temporal beat frequency equal to the difference frequency.

Interference Fringes

We will now add the additional restrictions that the two linear polarizations are parallel and that the two waves are at the same optical frequency. The expression for the intensity pattern now becomes

$$I(x, y, z) = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos[\Delta\phi(x, y, z)] \tag{16}$$

where $\Delta\phi = \phi_1 - \phi_2$ is the phase difference. This is the basic equation describing interference. The detected intensity varies cosinusoidally with the phase difference between the two waves as shown in Fig. 2. These alternating bright and dark bands in the

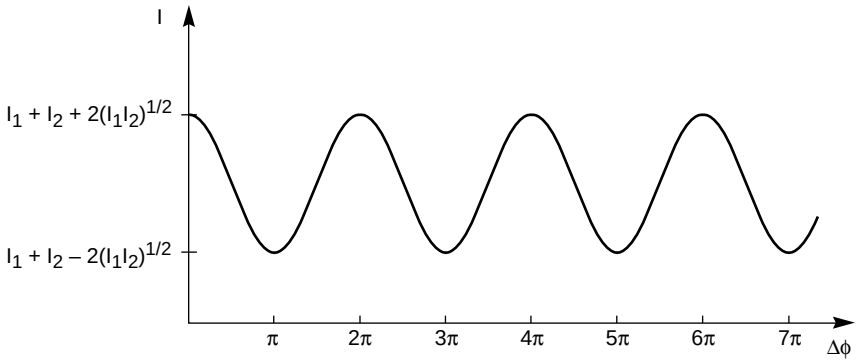


FIGURE 2 The variation in intensity as a function of the phase difference between two interfering waves.

TABLE 1 The Phase Difference and OPD for Bright and Dark Fringes (m an Integer)

	$\Delta\phi$	OPD
Bright fringe	$2m\pi$	$m\lambda$
Dark fringe	$2(m+1)\pi$	$(m+1/2)\lambda$

intensity pattern are referred to as *interference fringes,* and along a particular fringe, the phase difference is constant.

The phase difference is related to the difference in the optical path lengths between the source and the observation point for the two waves. This is the *optical path difference* (OPD):

$$\text{OPD} = \text{OPL}_1 - \text{OPL}_2 = \left(\frac{\lambda}{2\pi}\right)\Delta\phi \tag{17}$$

or

$$\Delta\phi = \left(\frac{2\pi}{\lambda}\right)\text{OPD} \tag{18}$$

The phase difference changes by 2π every time the OPD increases by a wavelength. The OPD is therefore constant along a fringe.

Constructive interference occurs when the two waves are in phase, and a bright fringe or maximum in the intensity pattern results. This corresponds to a phase difference of an integral number of 2π's or an OPD that is a multiple of the wavelength. A dark fringe or minimum in the intensity pattern results from *destructive interference* when the two waves are out of phase by π or the OPD is an odd number of half wavelengths. These results are summarized in Table 1. For conditions between these values, an intermediate value of the intensity results. Since both the OPD and the phase difference increase with the integer m, the absolute value of m is called the *order of interference.*

As we move from one bright fringe to an adjacent bright fringe, the phase difference changes by 2π. Each fringe period corresponds to a change in the OPD of a single wavelength. It is this inherent precision that makes interferometry such a valuable metrology tool. The wavelength of light is used as the unit of measurement. Interferometers can be configured to measure small variations in distance, index, or wavelength.

When two monochromatic waves are interfered, the interference fringes exist not only in the plane of observation, but throughout all space. This can easily be seen from Eq. (16) where the phase difference can be evaluated at any z position. In many cases, the observation of interference is confined to a plane, and this plane is usually assumed to be perpendicular to the z axis. The z dependence in Eq. (16) is therefore often not stated explicitly, but it is important to remember that interference effects will exist in other planes.

Fringe Visibility

It is often more convenient to rewrite Eq. (16) as

$$I(x, y) = I_0(x, y)\{1 + \gamma(x, y)\cos[\Delta\phi(x, y, z)]\} \tag{19}$$

or

$$I(x, y) = I_0(x, y)\{1 + \gamma(x, y)\cos[2\pi\text{OPD}(x, y)/\lambda]\} \tag{20}$$

where $I_0(x, y) = I_1(x, y) + I_2(x, y)$, and

$$\gamma(x, y) = \frac{2[I_1(x, y)I_2(x, y)]^{1/2}}{I_1(x, y) + I_2(x, y)} \tag{21}$$

Since the cosine averages to zero, $I_0(x, y)$ represents the average intensity, and $\gamma(x, y)$ is the local *fringe contrast* or *visibility*. The fringe visibility can also be equivalently calculated using the standard formula for modulation:

$$\gamma(x, y) = \frac{I_{\max}(x, y) - I_{\min}(x, y)}{I_{\max}(x, y) + I_{\min}(x, y)} \tag{22}$$

where $I_{\max}$ and $I_{\min}$ are the maximum and minimum intensities in the fringe pattern.

The fringe visibility will have a value between 0 and 1. The maximum visibility will occur when the two waves have equal intensity. Not surprisingly, the visibility will drop to zero when one of the waves has zero intensity. In general, the intensities of the two waves can vary with position, so that the average intensity and fringe visibility can also vary across the fringe pattern. The average intensity in the observation plane equals the sum of the individual intensities of the two interfering waves. The interference term redistributes this energy into bright and dark fringes.

Two Plane Waves

The first special case to consider is the interference of two plane waves of equal intensity, polarization and frequency. They are incident at angles θ_1 and θ_2 on the observation plane, as shown in Fig. 3. The plane of incidence is the x-z plane (the two k-vectors are contained in this plane). According to Eq. (5), the complex amplitude for each of these plane waves is

$$\mathbf{E}_i(x, y, z, t) = \mathbf{A}e^{i[\omega t - kz\cos(\theta_i) - kx\sin(\theta_i)]} \tag{23}$$

where the dot product has been evaluated. For simplicity we will place the observation plane at $z = 0$, and the phase difference between the two waves is

$$\Delta\phi(x, y) = kx(\sin\theta_1 - \sin\theta_2) = (2\pi x/\lambda)(\sin\theta_1 - \sin\theta_2) \tag{24}$$

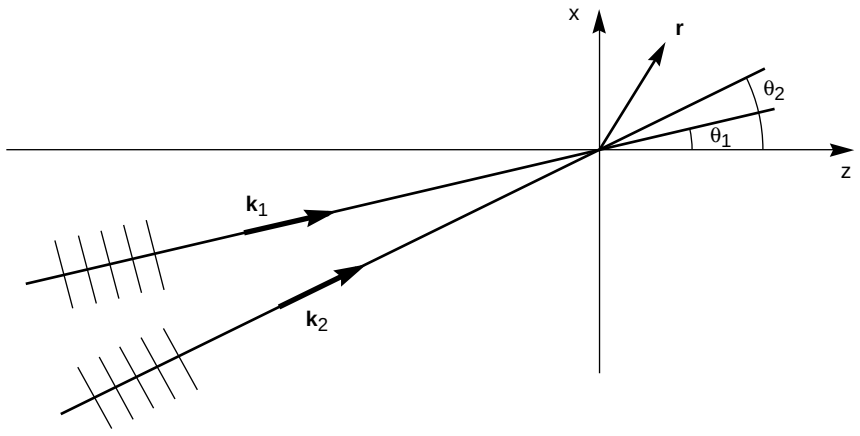


FIGURE 3 The geometry for the interference of two plane waves.

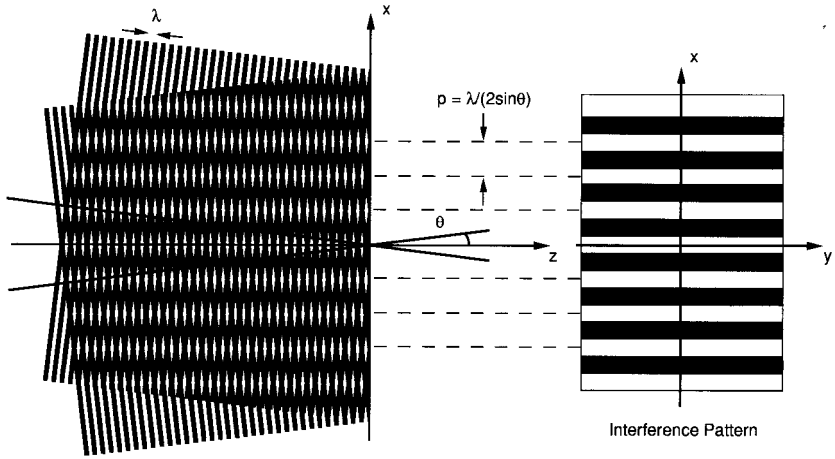


FIGURE 4 The interference of plane waves incident at $\pm\theta$ resulting in straight fringes.

The resulting intensity from Eq. (19) is

$$I(x, y) = I_0\{1 + \cos[(2\pi x/\lambda)(\sin\theta_1 - \sin\theta_2)]\} \tag{25}$$

where $I_0 = 2A^2$ is twice the intensity of each of the individual waves. Straight equispaced fringes are produced. The fringes are parallel to the y axis, and the fringe period depends on the angle between the two interfering beams.

The fringe period p is

$$p = \frac{\lambda}{\sin\theta_1 - \sin\theta_2} \tag{26}$$

and this result can also be obtained by noting that a bright fringe will occur whenever the phase difference equals a multiple of 2π. A typical situation for interference is that the two angles of incidence are equal and opposite, $\theta_1 = -\theta_2 = \theta$. The angle between the two beams is 2θ. Under this condition, the period is

$$p = \frac{\lambda}{2\sin\theta} \approx \frac{\lambda}{2\theta} \tag{27}$$

and the small-angle approximation is given. As the angle between the beams gets larger, the period decreases. For example, the period is 3.8λ at 15° (full angle of 30°) and is λ at 30° (full angle of 60°). The interference of two plane waves can be visualized by looking at the overlap or moiré of two wavefront patterns (Fig. 4). Whenever the lines representing the wavefronts overlap, a fringe will result. This description also clearly shows that the fringes extend parallel to the z axis and exist everywhere the two beams overlap.

Plane Wave and Spherical Wave

A second useful example to consider is the interference of a plane wave and a spherical wave. Once again the two waves have the same frequency. The plane wave is at normal incidence, the spherical wave is due to a source at the origin, and the observation plane is located at $z = R$. The wavefront shape at the observation plane will be a spherical shell of

radius R. Starting with Eq. (8), the complex amplitude of the spherical wave in the observation plane is

$$E(\rho, t) = (A/R)e^{i[\omega t - k(R^2+\rho^2)^{1/2}]} \approx (A/R)e^{i[\omega t - k(R+\rho^2/2R)]} \tag{28}$$

where $\rho = (x^2 + y^2)^{1/2}$, and the square root has been expanded in the second expression. This expansion approximates the spherical wave by a parabolic wave with the same vertex radius. An additional assumption is that the amplitude of the field A/R is constant over the region of interest. The field for the plane wave is found by evaluating Eq. (23) at $z = R$ and $\theta = 0$. The phase difference between the plane and the sphere is then

$$\Delta\phi(\rho) \approx \frac{\pi\rho^2}{\lambda R} \tag{29}$$

and the resulting intensity pattern is

$$I(\rho) = I_0\left[1 + \cos\left(\frac{\pi\rho^2}{\lambda R}\right)\right] \tag{30}$$

The fringe pattern is comprised of concentric circles, and the radial fringe spacing decreases as the radius ρ increases. The intensities of the two waves have been assumed to be equal at the observation plane. This result is valid only when ρ is much smaller than R.

The radius of the mth bright fringe can be found by setting $\Delta\phi = 2\pi m$:

$$\rho_m = \sqrt{2mR} \tag{31}$$

where m is an integer. The order of interference m increases with radius. Figure 5 shows a visualization of this situation using wavefronts. This fringe pattern is the Newton's ring pattern and is discussed in more detail later, under "Fizeau Interferometer." This picture also shows that the radii of the fringes increase as the square root of R.

The analysis of the spherical wave could also have been done by using the sag of a spherical wavefront to produce an OPD and then converting this value to a phase difference. The quadratic approximation for the sag of a spherical surface is $\rho^2/2R$. This

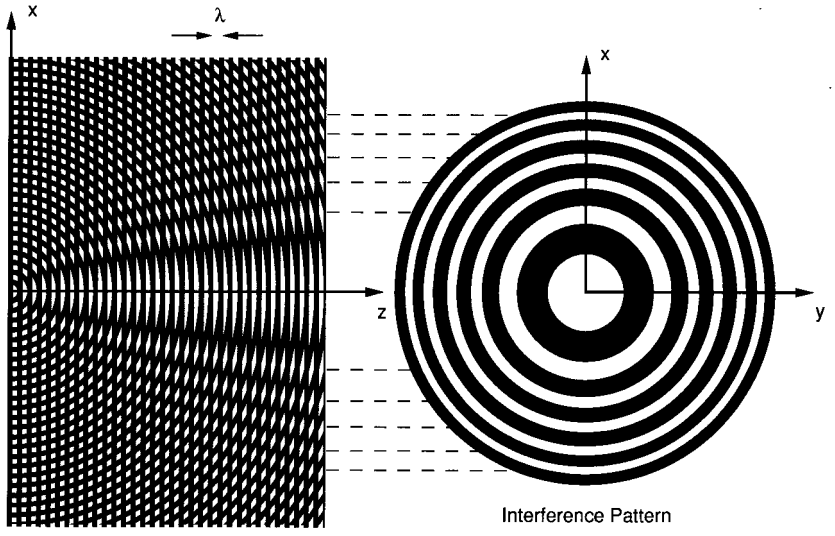


FIGURE 5 The interference of a plane wave and a spherical wave.

corresponds to the OPD between the spherical and planar wavefronts. The equivalent phase difference [Eq. (18)] is then $\pi\rho^2/\lambda R$, as before.

Two Spherical Waves

When considering two spherical waves, there are two particular geometries that we want to examine. The first places the observation plane perpendicular to a line connecting the two sources, and the second has the observation plane parallel to this line. Once again, the sources are at the same frequency.

When the observations are made on a plane perpendicular to a line connecting the two sources, we can use Eq. (28) to determine the complex amplitude of the two waves:

$$E_i(\rho, t) \approx (A/R)e^{i[\omega t - k(R_i + \rho^2/2R_i)]} \tag{32}$$

Let $d = R_1 - R_2$ be the separation of the two sources. For simplicity, we have also assumed that the amplitudes of the two waves are equal (R is an average distance). The phase difference between the two waves is

$$\Delta\phi = \left(\frac{\pi\rho^2}{\lambda}\right)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) + \frac{2\pi d}{\lambda} \approx \frac{2\pi d}{\lambda} - \left(\frac{\pi\rho^2}{\lambda}\right)\left(\frac{d}{R^2}\right), \tag{33}$$

where the approximation $R_1R_2 \approx R^2$ has been made. There are two terms to this phase difference. The second is a quadratic phase term identical in form to the result obtained from spherical and plane waves. The pattern will be symmetric around the line connecting the two sources, and its appearance will be similar to Newton's rings. The equivalent radius of the spherical wave in Eq. (29) is R^2/d. The first term is a constant phase shift related to the separation of the two sources. If this term is not a multiple of 2π, the center of the fringe pattern will not be a bright fringe; if the term is π, the center of the pattern will be dark. Except for the additional phase shift, this intensity pattern is not distinguishable from the result in the previous section. It should be noted, however, that a relative phase shift can be introduced between a spherical wave and a plane wave to obtain this same result.

An important difference between this pattern and the Newton's ring pattern is that the order of interference ($|m|$ defined by $\Delta\phi = 2\pi m$) or phase difference is a maximum at the center of the pattern and decreases with radius. The phase difference is zero at some finite radius. The Newton's ring pattern formed between a plane and a spherical wave has a minimum order of interference at the center of the pattern. This distinction is important when using polychromatic sources.

There are several ways to analyze the pattern that is produced on a plane that is parallel to a line connecting the two sources. We could evaluate the complex amplitudes by using Eq. (28) and moving the center of the spherical waves to $\pm d/2$ for the two sources. An equivalent method is to compare the wavefronts at the observation plane. This is shown in Fig. 6. The OPD between the two wavefronts is

$$\mathrm{OPD}(x, y) = \frac{[(x + d/2)^2 + y^2]}{2L} - \frac{[(x - d/2)^2 + y^2]}{2L} \tag{34}$$

where the quadratic approximation for the wavefront sag has been assumed, and L is the distance between the sources and the observation plane. After simplification, the OPD and phase differences are

$$\mathrm{OPD}(x, y) = \frac{xd}{L} \tag{35}$$

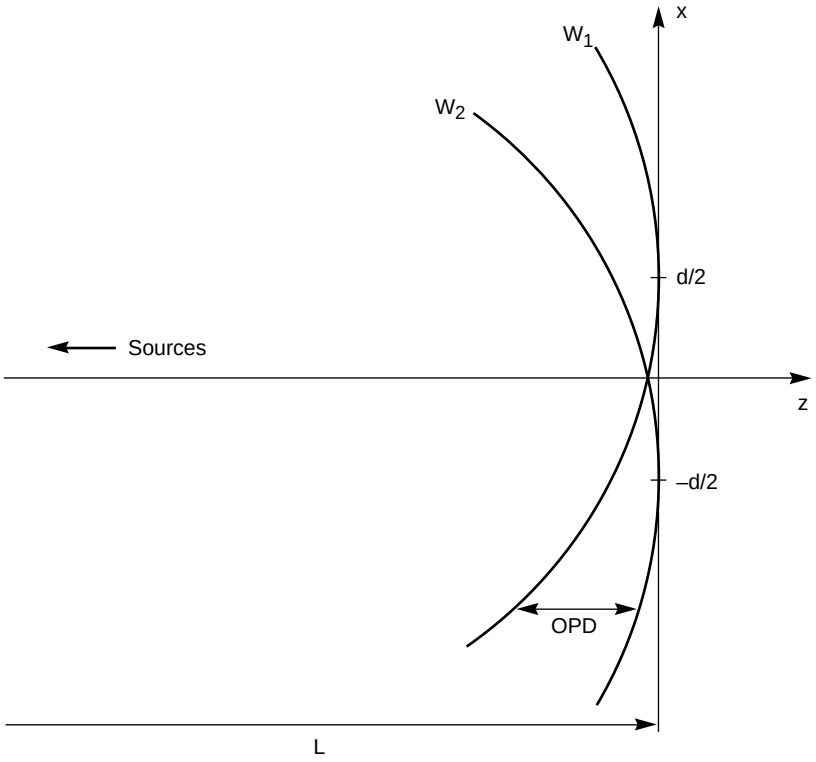


FIGURE 6 The interference of two spherical waves on a plane parallel to the sources.

and

$$\Delta\phi(x, y) = \frac{2\pi x d}{\lambda L} \tag{36}$$

Straight equispaced fringes parallel to the y axis are produced. The period of the fringes is $\lambda L/d$. This fringe pattern is the same as that produced by two plane waves. Note that these fringes increase in spacing as the distance from the sources increases. The approximations used require that L be much larger than ρ and d.

Figure 7 shows the creation of the fringe patterns for two point sources. The full three-dimensional pattern is a series of nested hyperboloids symmetric about the line connecting the sources. Above the two sources, circular fringes approximating Newton's rings are produced, and perpendicular to the sources, the fringes appear to be straight and equispaced.

Aberrated Wavefronts

When an aberrated or irregularly shaped wavefront is interfered with a reference wavefront, an irregularly shaped fringe pattern is produced. However, the rules for analyzing this pattern are the same as with any two wavefronts. A given fringe represents a contour of constant OPD or phase difference between the two wavefronts. Adjacent fringes differ in OPD by one wavelength or equivalently correspond to a phase difference

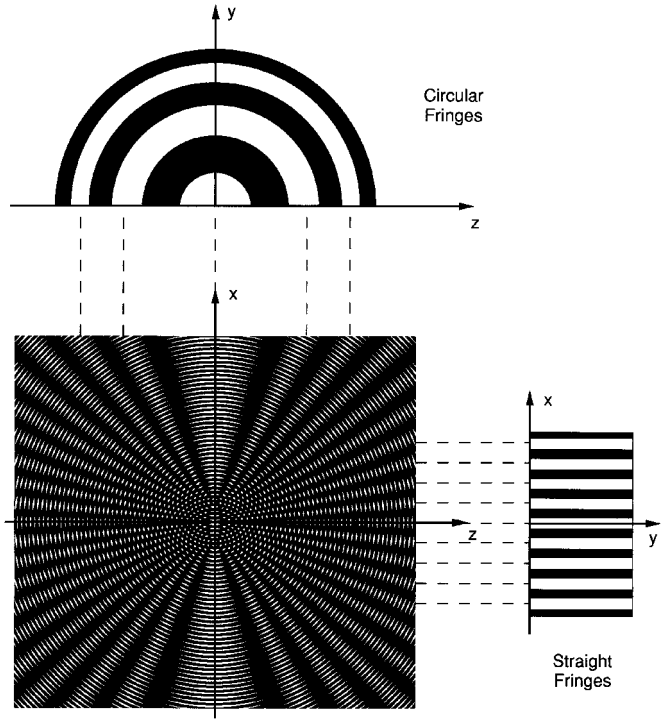


FIGURE 7 The interference of two spherical waves.

of 2π. If the reference is a plane wave, the absolute shape of the irregular wavefront is obtained. If the reference is a spherical wave, or another aberrated wave, the measured OPD or phase difference map represents the difference between the two wavefronts.

Temporal Beats

In Eq. (15) it was noted that if the waves are at different frequencies, the interference effects are modulated by a beat frequency. Rewriting this expression assuming equal-intensity parallel-polarized beams produces

$$I(x, y, t) = I_0\{1 + \cos [2\pi\, \Delta\nu t - \Delta\phi(x, y)]\} \tag{37}$$

where $\Delta\nu = \nu_1 - \nu_2$. The intensity at a given location will now vary sinusoidally with time at the beat frequency $\Delta\nu$. The phase difference $\Delta\phi$ appears as a spatially varying phase shift of the beat frequency. This is the basis of the heterodyne technique used in a number of interferometers. It is commonly used in distance-measuring interferometers.

In order for a heterodyne system to work, there must be a phase relationship between the two sources even though they are at different frequencies. One common method for obtaining this is accomplished by starting with a single source, splitting it into two beams, and frequency-shifting one beam with a known Doppler shift. The system will also work in

reverse; measure the interferometric beat frequency to determine the velocity of the object producing the Doppler shift.

Coherence

Throughout this discussion of fringe patterns, we have assumed that the two sources producing the two waves have the same frequency. In practice, this requires that both sources be derived from a single source. Even when two different frequencies are used [Eq. (37)] there must be an absolute phase relation between the two sources. If the source has finite size, it is considered to be composed of a number of spatially separated, independently radiating point sources. If the source has a finite spectral bandwidth, it is considered to be composed of a number of spatially coincident point sources with different frequencies. These reductions in the spatial or temporal coherence of the source will decrease the visibility of the fringes at different locations in space. This is referred to as *fringe localization.* These effects will be discussed later in this chapter and also in Chap. 4, "Coherence Theory."

There are two general methods to produce mutually coherent waves for interference. The first is called *wavefront division,* where different points on a wavefront are sampled to produce two new wavefronts. The second is *amplitude division,* where some sort of beamsplitter is used to divide the wavefront at a given location into two separate wavefronts. These methods are discussed in the next sections.

2.5 INTERFERENCE BY WAVEFRONT DIVISION

Along a given wavefront produced by a monochromatic point source, the wavefront phase is constant. If two parts of this wavefront are selected and then redirected to a common volume in space, interference will result. This is the basis for *interference by wavefront division.*

Young's Double-Slit Experiment

In 1801, Thomas Young performed a fundamental experiment for demonstrating interference and the wave nature of light. Monochromatic light from a single pinhole illuminates an opaque screen with two additional pinholes or slits. The light diffracts from these pinholes and illuminates a viewing screen at a distance large compared to the pinhole separation. Since the light illuminating the two pinholes comes from a single source, the two diffracted wavefronts are coherent and interference fringes form where the beams overlap.

In the area where the two diffracted beams overlap, they can be modeled as two spherical waves from two point sources, and we already know the form of the solution for the interference from our earlier discussion. Equispaced straight fringes are produced, and the period of the fringes is $\lambda L/d$, where L is the distance to the screen and d is the separation of the pinholes. The fringes are oriented perpendicular to the line connecting the two pinholes.

Even though we already know the answer, there is a classic geometric construction we should consider that easily gives the OPD between the two wavefronts at the viewing

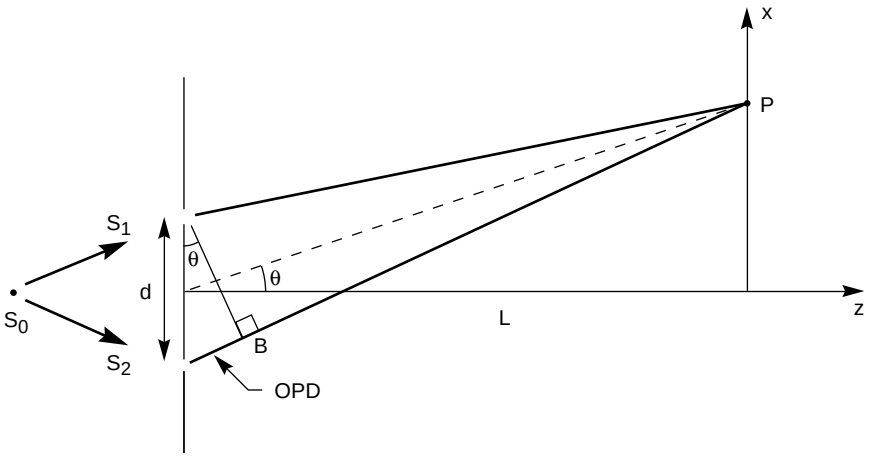


FIGURE 8 Young's double-slit experiment.

screen. This is shown in Fig. 8. S_0 illuminates both S_1 and S_2 and is equidistant from both slits. The OPD for an observation point P at an angle θ or position x is

$$\text{OPD} = \overline{S_2P} - \overline{S_1P} \tag{38}$$

We now draw a line from S_1 to B that is perpendicular to the second ray. Since L is much larger than d, the distances from B to P and S_1 to P are approximately equal. The OPD is then

$$\text{OPD} \approx \overline{S_2B} = d \sin\theta \approx d\theta \approx \frac{dx}{L} \tag{39}$$

and constructive interference or a bright fringe occurs when the OPD is a multiple of the wavelength: $\text{OPD} = m\lambda$, where m is an integer. The condition for the mth order bright fringe is

$$\text{Bright fringe:} \quad \sin(\theta) \approx \theta = \frac{m\lambda}{d} \qquad \text{or} \qquad x = \frac{m\lambda L}{d} \tag{40}$$

This construction is useful not only for interference situations, but also for diffraction analysis.

Effect of Slit Width

The light used to produce the interference pattern is diffracted by the pinholes or slits. Interference is possible only if light is directed in that direction. The overall interference intensity pattern is therefore modulated by the single-slit diffraction pattern (assuming slit apertures):

$$I(x) = I_0 \,\text{sinc}^2\left(\frac{Dx}{\lambda L}\right)\left[1 + \gamma(x)\cos\left(\frac{2\pi x d}{\lambda L}\right)\right] \tag{41}$$

where D is the slit width, and a one-dimensional expression is shown. The definition of a sinc function is

$$\text{sinc}(\alpha) = \frac{\sin(\pi\alpha)}{\pi\alpha} \tag{42}$$

where the zeros of the function occur when the argument α is an integer. The intensity variation in the y direction is due to diffraction only and is not shown. Since the two slits are assumed to be illuminated by a single source, there are no coherence effects introduced by using a pinhole or slit of finite size.

The term $\gamma(x)$ is included in Eq. (41) to account for variations in the fringe visibility. These could be due to unequal illumination of the two slits, a phase difference of the light reaching the slits, or a lack of temporal or spatial coherence of the source S_0.

Other Arrangements

Several other arrangements for producing interference by division of wavefront are shown in Fig. 9. They all use a single source and additional optical elements to produce two separate and mutually coherent sources. Fresnel's biprism and mirror produce the two

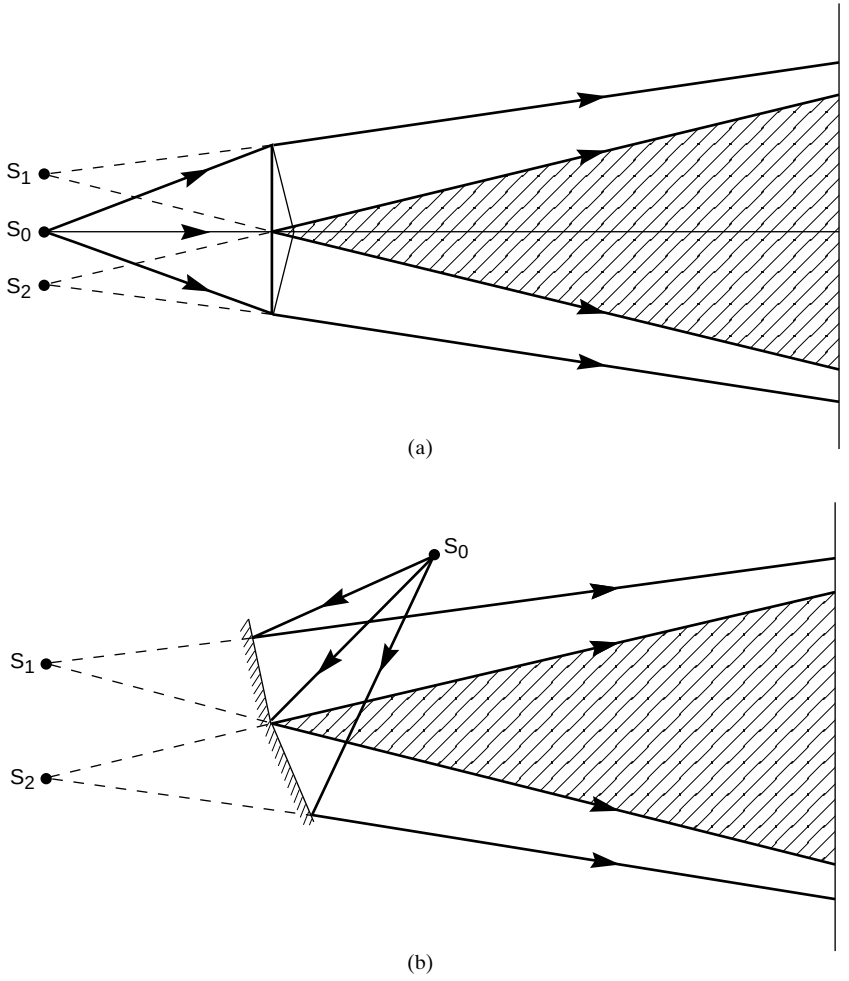
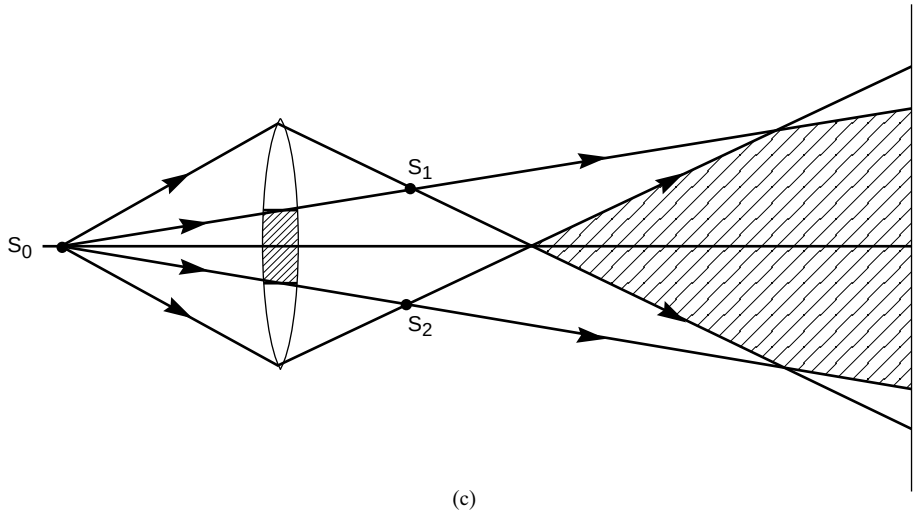
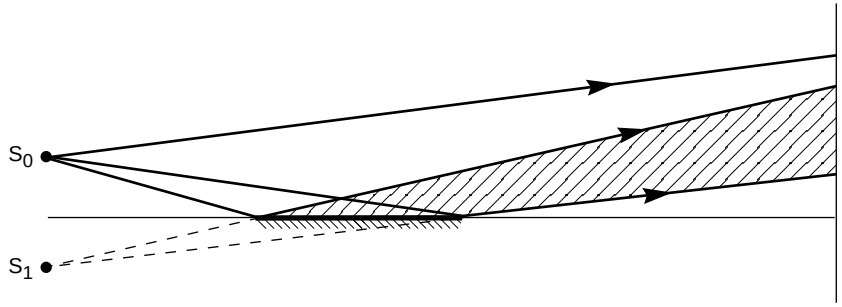


FIGURE 9 Arrangements for interference by division of wavefront: (*a*) Fresnel's biprism; (*b*) Fresnel's mirror; (*c*) Billet's split lens; and (*d*) Lloyd's mirror.



(c)



(d)

FIGURE 9 (*Continued*.)

virtual source images, Billet's split lens produces two real source images, and Lloyd's mirror produces a single virtual source image as a companion to the original source. Interference fringes form wherever the two resulting waves overlap (shaded regions). One significant difference between these arrangements and Young's two slits is that a large section of the initial wavefront is used instead of just two points. All of these systems are much more light efficient, and they do not rely on diffraction to produce the secondary wavefronts.

In the first three of these systems, a bright fringe is formed at the zero OPD point between the two sources as in the double-slit experiment. With Lloyd's mirror, however, the zero OPD point has a dark fringe. This is due to the π phase shift that is introduced into one of the beams on reflection from the mirror.

Source Spectrum

The simple fringe pattern produced by the two-slit experiment provides a good example to examine the effects of a source with a finite spectrum. In this model, the source can be considered to be a collection of sources, each radiating independently and at a different wavelength. All of these sources are colocated to produce a point source. (Note that this is

an approximation, as a true point source must be monochromatic.) At each wavelength, an independent intensity pattern is produced:

$$I(x, \lambda) = I_0\left[1 + \cos\left(\frac{2\pi x d}{\lambda L}\right)\right] = I_0\left[1 + \cos\left(\frac{2\pi\ \mathrm{OPD}}{\lambda}\right)\right] \tag{43}$$

where the period of the fringes is $\lambda L/d$, and a fringe visibility of one is assumed. The total intensity pattern is the sum of the individual fringe patterns:

$$I(x) = \int_0^\infty S(\lambda) I(x, \lambda)\, d\lambda = \int_0^\infty S(\nu) I(x, \nu)\, d\nu \tag{44}$$

where $S(\lambda)$ or $S(\nu)$ is the source intensity spectrum which serves as a weighting function.

The effect of this integration can be seen by looking at a simple example where the source is composed of three different wavelengths of equal intensity. To further aid in visualization, let's use Blue (400 nm), Green (500 nm), and Red (600 nm). The result is shown in Fig. 10*a*. There are three cosine patterns, each with a period proportional to the wavelength. The total intensity is the sum of these curves. All three curves line up when the OPD is zero ($x = 0$), and the central bright fringe is now surrounded by two-colored dark fringes. These first dark fringes have a red to blue coloration with increasing OPD. As we get further away from the zero OPD condition, the three patterns get out of phase, the pattern washes out, and the color saturation decreases. This is especially true when the source is composed of more than three wavelengths.

It is common in white light interference situations for one of the two beams to undergo an additional π phase shift. This is the situation in Lloyd's mirror. In this case, there is a central dark fringe at zero OPD with colored bright fringes on both sides. This is shown in Fig. 10*b*, and the pattern is complementary to the previous pattern. In this case the first

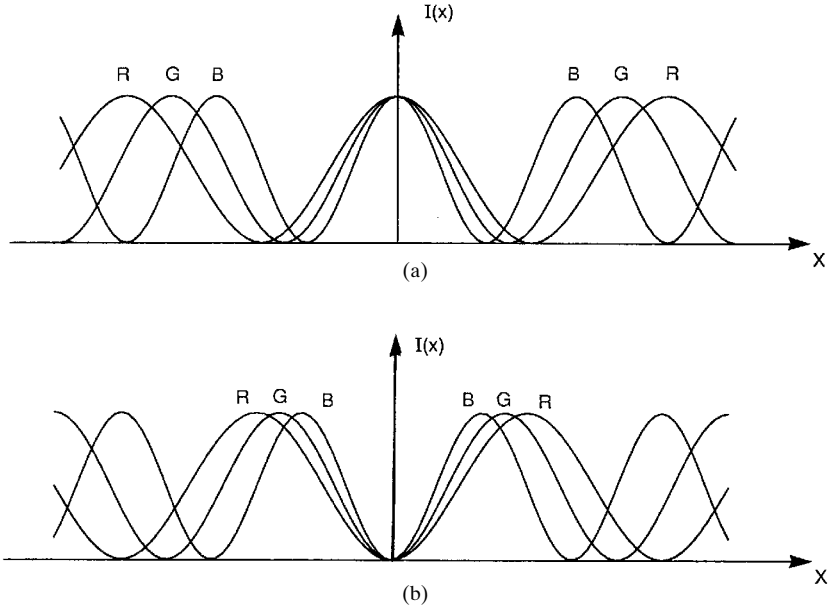


FIGURE 10 The interference pattern produced by a source with three separate wavelengths: (*a*) zero OPD produces a bright fringe; and (*b*) zero OPD produces a dark fringe.

bright fringe shows a blue to red color smear. The dark central fringe is useful in determining the location of zero OPD between the two beams.

The overall intensity pattern and resulting fringe visibility can be computed for a source with a uniform frequency distribution over a frequency range of $\Delta\nu$:

$$I(x) = \frac{1}{\Delta\nu}\int_{\nu_0-\Delta\nu/2}^{\nu_0+\Delta\nu/2} I(x,\,\nu)\,d\nu = \frac{1}{\Delta\nu}\int_{\nu_0-\Delta\nu/2}^{\nu_0+\Delta\nu/2} I_0\left[1+\cos\left(\frac{2\pi\nu x d}{cL}\right)\right]d\nu \tag{45}$$

where ν_0 is the central frequency, and the $1/\Delta\nu$ term is a normalization factor to assure that the average intensity is I_0. After integration and simplification, the result is

$$I(x) = I_0\left[1+\operatorname{sinc}\left(\frac{xd\,\Delta\nu}{cL}\right)\cos\left(\frac{2\pi\nu_0 x d}{cL}\right)\right] \tag{46}$$

where the sinc function is defined in Eq. (42). A fringe pattern due to the average optical frequency results, but it is modulated by a sinc function that depends on $\Delta\nu$ and x. The absolute value of the sinc function is the fringe visibility $\gamma(x)$, and it depends on both the spectral width and position of observation. The negative portions of the sinc function correspond to a π phase shift of the fringes.

It is informative to rewrite this expression in terms of the OPD:

$$I(x) = I_0\left[1+\operatorname{sinc}\left(\frac{\mathrm{OPD}\,\Delta\nu}{c}\right)\cos\left(\frac{2\pi\,\mathrm{OPD}}{\lambda_0}\right)\right] \tag{47}$$

where λ_0 is the wavelength corresponding to ν_0. Good fringe visibility is obtained only when either the spectral width is small (the source is quasi-monochromatic) or the OPD is small. The fringes are *localized* in certain areas of space. This result is consistent with the earlier graphical representations. In the area where the OPD is small, the fringes are in phase for all wavelengths. As the OPD increases, the fringes go out of phase since they all have different periods, and the intensity pattern washes out.

This result turns out to be very general: for an incoherent source, the fringes will be localized in the vicinity of zero OPD. There are two other things we should notice about this result. The first is that the first zero of the visibility function occurs when the OPD equals $c/\Delta\nu$. This distance is known as the *coherence length* as it is the path difference over which we can obtain interference. The second item is that the visibility function is a scaled version of the Fourier transform of the source frequency spectrum. It is evaluated for the OPD at the measurement location. The Fourier transform of a uniform distribution is a sinc function. We will discuss this under ''Coherence and Interference'' later in the chapter.

2.6 INTERFERENCE BY AMPLITUDE DIVISION

The second general method for producing interference is to use the same section of a wavefront from a single source for both resulting wavefronts. The original wavefront amplitude is split into two or more parts, and each fraction is directed along a different optical path. These waves are then recombined to produce interference. This method is called *interference by amplitude division.* There are a great many interferometer designs based on this method. A few will be examined here, and many more will be discussed in Chap. 21 of Vol. 2, ''Interferometers.''

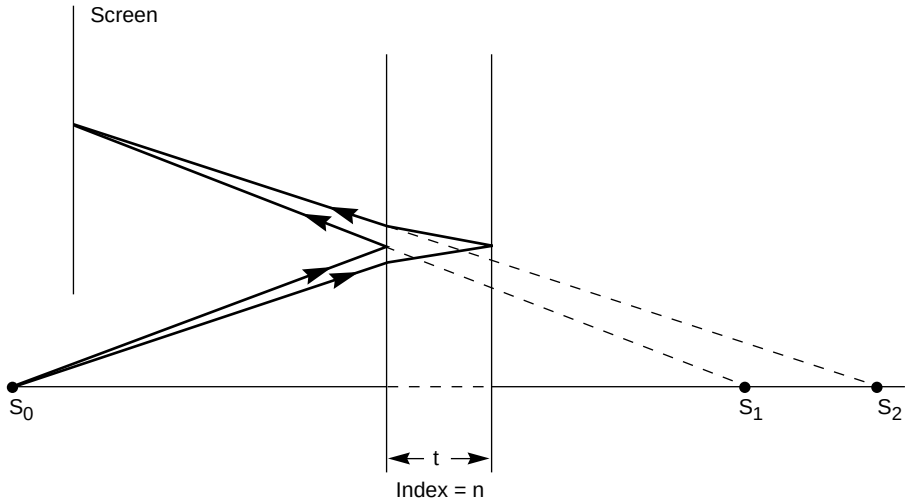


FIGURE 11 Interference from a plane-parallel plate and a point source.

Plane-Parallel Plate

A first example of interference by amplitude division is a plane-parallel plate illuminated by a monochromatic point source. Two virtual images of the point source are formed by the Fresnel reflections at the two surfaces, as shown in Fig. 11. Associated with each of the virtual images is a spherical wave, and interference fringes form wherever these two waves overlap. In this case, this is the volume of space on the source side of the plate. The pattern produced is the same as that found for the interference of two spherical waves (discussed earlier under ''Two Spherical Waves''), and nonlocalized fringes are produced. The pattern is symmetric around the line perpendicular to the plate through the source. If a screen is placed along this axis, a pattern similar to circular Newton's ring fringes are produced as described by Eq. (33), where $d = 2t/n$ is now the separation of the virtual sources. The thickness of the plate is t, its index is n, and the distance R is approximately the screen-plate separation plus the source-plate separation. We have ignored multiple reflections in the plate. As with the interference of two spherical waves, the order of interference is a maximum at the center of the pattern.

The interference of two plane waves can be obtained by illuminating a wedged glass plate with a plane wavefront. If the angle of incidence on the first surface is θ and the wedge angle is α, two plane waves are produced at angles θ and $\theta + 2n\alpha$ due to reflections at the front and rear surfaces. Straight equispaced fringes will result in the volume of space where the two reflected waves overlap. The period of these fringes on a screen parallel to the plate is given by Eq. (26), where the two reflected angles are used.

Extended Source

An *extended source* is modeled as a collection of independent point sources. If the source is quasi-monochromatic, all of the point sources radiate at the same nominal frequency, but without a phase relationship. Each point source will produce its own interference pattern, and the net intensity pattern is the sum or integral of all the individual intensity patterns. This is the spatial analogy to the temporal average examined earlier under ''Source Spectrum.''

With an extended source, the fringes will be localized where the individual fringe position or spacing is not affected by the location of the point sources that comprise the extended source. We know from our previous examples that a bright fringe (or a dark fringe, depending on phase shifts) will occur when the OPD is zero. If there is a location where the OPD is zero independent of source location, all of the individual interference patterns will be in phase, and the net pattern will show good visibility. In fact, the three-dimensional fringe pattern due to a point source will tend to shift or pivot around this zero-OPD location as the point source location is changed. The individual patterns will therefore be out of phase in areas where the OPD is large, and the average intensity pattern will tend to wash out in these regions as the source size increases.

The general rule for fringe visibility with an extended quasi-monochromatic source is that the fringes will be localized in the region where the OPD between the two interfering wavefronts is small. For a wedged glass plate, the fringes are localized in or near the wedge, and the best visibility occurs as the wedge thickness approaches zero and is perhaps just a few wavelengths. The allowable OPD will depend on the source size and the method of viewing the fringes. This result explains why, under natural light, interference effects are seen in thin soap bubbles but not with other thicker glass objects. An important exception to this rule is the plane-parallel plate where the fringes are localized at infinity.

Fringes of Equal Inclination

There is no section of a plane-parallel plate that produces two reflected wavefronts with zero OPD. The OPD is constant, and we would expect, based on the previous section, that no high-visibility fringes would result with an extended source. If, however, a lens is used to collect the light reflected from the plate, fringes are formed in the back focal plane of the lens. This situation is shown in Fig. 12, and any ray leaving the surface at a particular angle θ is focused to the same point P. For each incident ray at this angle, there are two parallel reflected rays: one from the front surface and one from the back surface. The reflections from different locations on the plate at this angle are due to light from different points in the extended source. The OPD for any pair of these reflected rays is the same regardless of the source location. These rays will interfere at P and will all have the same phase difference. High-visibility fringes result. Different points in the image plane correspond to different angles. The formation of these fringes localized at infinity depends on the two surfaces of the plate being parallel.

The OPD between the reflected rays is a function of the angle of incidence θ, the plate index n, and thickness t:

$$\text{OPD} = 2nt \cos \theta' \tag{48}$$

where θ' is the internal angle. Taking into account the half-wave shift due to the phase change difference of π between an internal and an external reflection, a dark fringe will result for angles satisfying

$$2nt \cos \theta' = m\lambda \qquad \text{or} \qquad \cos \theta' = \frac{m\lambda}{2nt} \tag{49}$$

where m is an integer. Since only the angle of incidence determines the properties of the interference (everything else is constant), these fringes are called *fringes of equal inclination.* They appear in the back focal plane of the lens and are therefore localized at infinity since infinity is conjugate to the focal plane. As the observation plane is moved away from the focal plane, the visibility of the fringes will quickly decrease.

When the axis of the lens is normal to the surfaces of the plate, a beamsplitter arrangement is required to allow light from the extended source to be reflected into the lens as shown in Fig. 13. Along the axis, $\theta = \theta' = 90°$, and symmetry requires that the

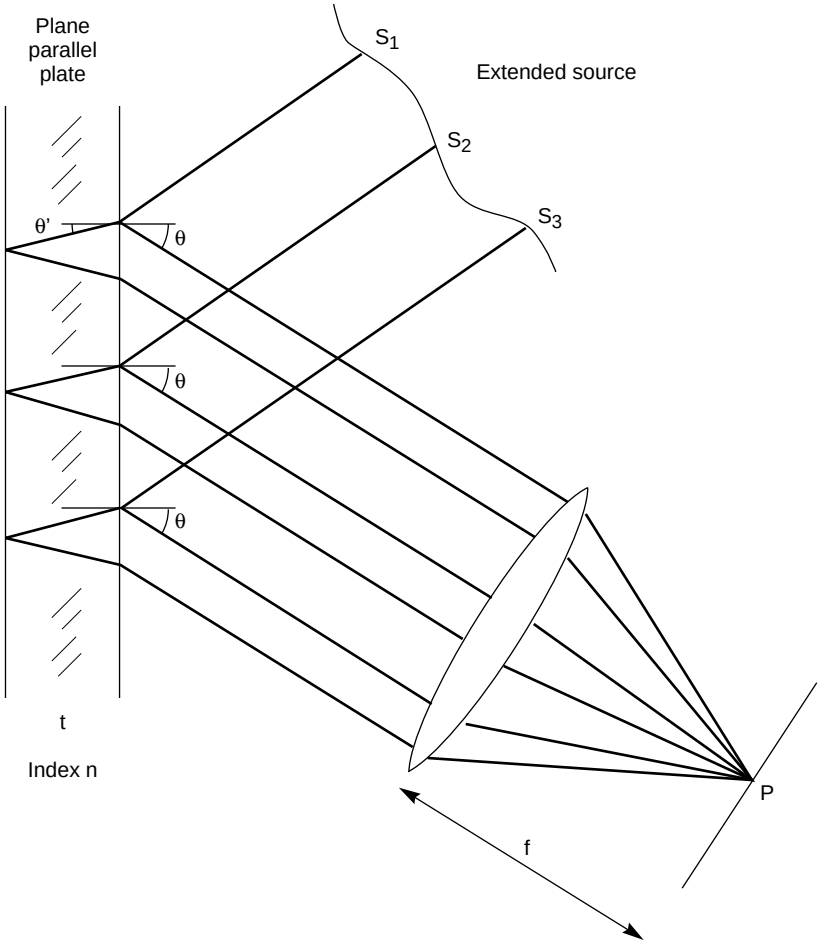


FIGURE 12 The formation of fringes of equal inclination.

fringes are concentric about the axis. In this special case, these fringes are called *Haidinger fringes,* and they are identical in appearance to Newton's rings [Eq. (30)]. If there is an intensity maximum at the center, the radii of the other bright fringes are proportional to the square roots of integers. As with other fringes formed by a plane-parallel plate (discussed earlier), the order of interference decreases with the observation radius on the screen. As θ' increases, the value of m decreases.

Fringes of Equal Thickness

The existence of fringes of equal inclination depends on the incident light being reflected by two parallel surfaces, and the angle of incidence is the mechanism which generates changes in the OPD. There are many arrangements with an extended source where the reflections are not parallel, and the resulting changes in OPD dominate the angle-of-incidence considerations. The fringes produced in this situation are called *fringes of equal*

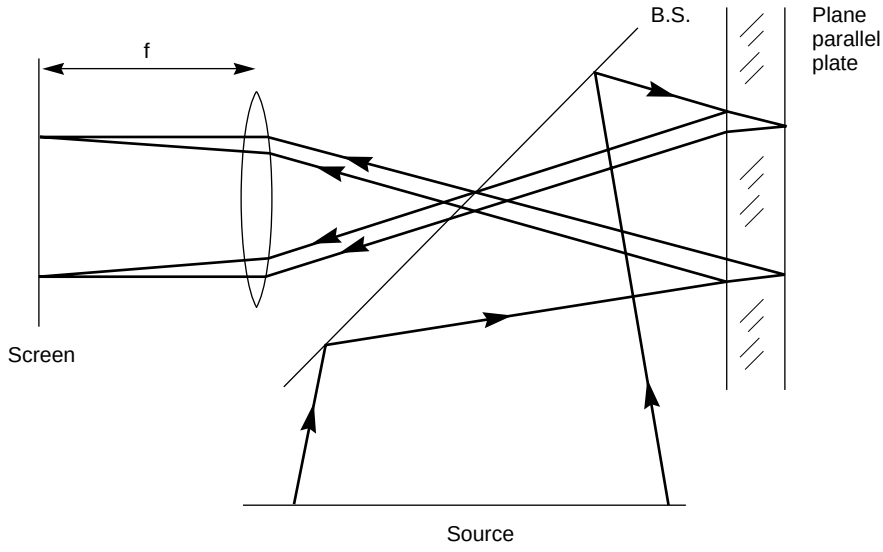


FIGURE 13 The formation of Haidinger fringes.

thickness, and we have stated earlier that they will be localized in regions where the OPD between the two reflections is small.

An example of fringes of equal thickness occurs with a wedged glass plate illuminated by a quasi-monochromatic extended source. We know that for each point in the source, a pattern comprised of equispaced parallel fringes results, and the net pattern is the sum of all of these individual patterns. However, it is easier to examine this summation by looking at the OPD between the two reflected rays reaching an observation point *P* from a source point *S.* This is shown in Fig. 14. The wedge angle is α, the thickness of the plate at this

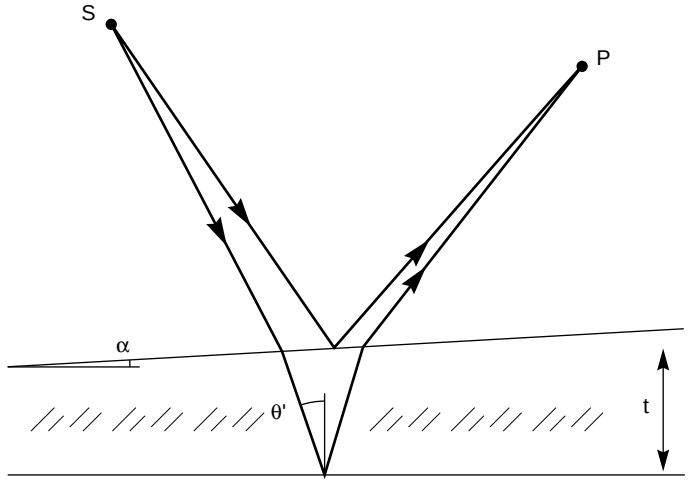


FIGURE 14 The ray path between a point source and an observation point for a wedged plate.

location is t, its index is n, and the internal ray angle is θ'. The exact OPD is difficult to calculate, but under the assumption that α is small and the wedge is sufficiently thin, the following result for the OPD is obtained:

$$\text{OPD} \approx 2nt \cos \theta' \tag{50}$$

As other points on the source are examined, the reflection needed to get light to the observation point will move to a different location on the plate, and different values of both t and θ' will result. Different source points may have greatly different OPDs, and in general the fringe pattern will wash out in the vicinity of P.

This reduction in visibility can be avoided if the observation point is placed in or near the wedge. In this case, all of the paths between S and P must reflect from approximately the same location on the wedge, and the variations in the thickness t are essentially eliminated. The point P where the two reflected rays cross may be virtual. The remaining variations in the OPD are from the different θ's associated with different source points. This variation may be limited by observing the fringe pattern with an optical system having a small entrance pupil. This essentially limits the amount of the source that is used to examine any area on the surface. A microscope or the eye focused on the wedge can be used to limit the angles. If the range of values of θ' is small, high-visibility fringes will appear to be localized at the wedge. The visibility of the fringes will decrease as the wedge thickness increases.

It is common to arrange the system so that the fringes are observed in a direction approximately normal to the surface. Taking into account the additional phase shift introduced at the reflection from one of the surfaces, the conditions for bright and dark fringes are then

$$\text{Bright:} \quad 2nt - \frac{\lambda}{2} = m\lambda \tag{51}$$

and

$$\text{Dark:} \quad 2nt = m\lambda \tag{52}$$

where m is an integer greater than or equal to zero. Since t increases linearly across the wedge, the observed pattern will be straight equispaced fringes.

These same conditions hold for any plate where the two surfaces are not parallel. The surfaces may have any shape, and as long as the surface angles are small and the plate is relatively thin, high-visibility fringes localized in the plate are observed. Along a given fringe the value of m is constant, so that a fringe represents a contour of constant optical path length nt. If the index is constant, we have fringes of equal thickness. The fringes provide a contour map of the plate thickness, and adjacent fringes correspond to a change of thickness of $\lambda/2n$. An irregularly shaped pattern will result from the examination of a plate of irregular thickness.

Thin Films

With the preceding background, we can easily explain the interference characteristics of *thin films*. There are two distinct types of films to consider. The first is a thin film of nonuniform thickness, and examples are soap bubbles and oil films on water. The second type is a uniform film, such as would be obtained by vacuum deposition and perhaps used as an antireflection coating. Both of these films share the characteristic of being extremely thin—usually not more than a few wavelengths thick and often just a fraction of a wavelength thick.

With a nonuniform film, fringes of equal thickness localized in the film are produced.

There will be a dark fringe in regions of the film where it is substantially thinner than a half wave. We are assuming that the film is surrounded by a lower-index medium such as air so that there is an extra π phase shift. If white light is used for illumination, colored bands will be produced similar to those diagramed in Fig. 10*b* (the curves would need to be modified to rescale the x axis to OPD or film thickness). Each color will produce its first maximum in intensity when the optical thickness of the film is a quarter of that wavelength. As the film thickness increases, the apparent fringe color will first be blue, then green, and finally red. These colored fringes are possible because the film is very thin, and the order of interference m is often zero or one [Eqs. (51) and (52)]. The interference patterns in the various colors are just starting to get out of phase, and interference colors are visible. As the film thickness increases, the various wavelength fringes become jumbled, and distinct fringe patterns are no longer visible.

When a uniform thin film is examined with an extended source, fringes of equal inclination localized at infinity are produced. These fringes will be very broad since the thickness of the film is very small, and large angles will be required to obtain the necessary OPD for a fringe [Eq. (49)]. A common use of this type of film is as an antireflection coating. In this application, a uniform coating that has an optical thickness of a quarter wavelength is applied to a substrate. The coating index is lower than the substrate index, so an extra phase shift is not introduced. A wave at normal incidence is reflected by both surfaces of the coating, and these reflected waves are interfered. If the incident wavelength matches the design of the film, the two reflected waves are out of phase and interfere destructively. The reflected intensity will depend on the Fresnel reflection coefficients at the two surfaces, but will be less than that of the uncoated surface. When a different wavelength is used or the angle of incidence is changed, the effectiveness of the antireflection coating is reduced. More complicated film structures comprised of many layers can be produced to modify the reflection or transmission characteristics of the film.

Fizeau Interferometer

The *Fizeau interferometer* compares one optical surface to another by placing them in close proximity. A typical arrangement is shown in Fig. 15, where the extended source is filtered to be quasi-monochromatic. A small air gap is formed between the two optical surfaces, and fringes of equal thickness are observed between the two surfaces. Equations (51) and (52) describe the location of the fringes, and the index of the thin wedge is now that of air. Along a fringe, the gap is of constant thickness, and adjacent fringes correspond to a change of thickness of a half wavelength. This interferometer is sometimes referred to as a *Newton interferometer.*

This type of interferometer is the standard test instrument in an optical fabrication shop. One of the two surfaces is a reference or known surface, and the interferometric comparison of this reference surface and the test surface shows imperfections in the test part. Differences in radii of the two surfaces are also apparent. The fringes are easy to interpret, and differences of as little as a twentieth of a wavelength can be visually measured. These patterns and this interferometer are further discussed in Chap. 30, Vol. 2, "Optical Testing." The interferometer is often used without the beamsplitter, and the fringes are observed in the direct reflection of the source from the parts.

The classic fringe pattern produced by a Fizeau interferometer is *Newton's rings.* These are obtained by comparing a convex sphere to a flat surface. The parabolic approximation for the sag of a sphere of radius R is

$$\text{sag}\,(\rho) = \frac{\rho^2}{2R} \tag{53}$$

and ρ is the radial distance from the vertex of the sphere. If we assume the two surfaces

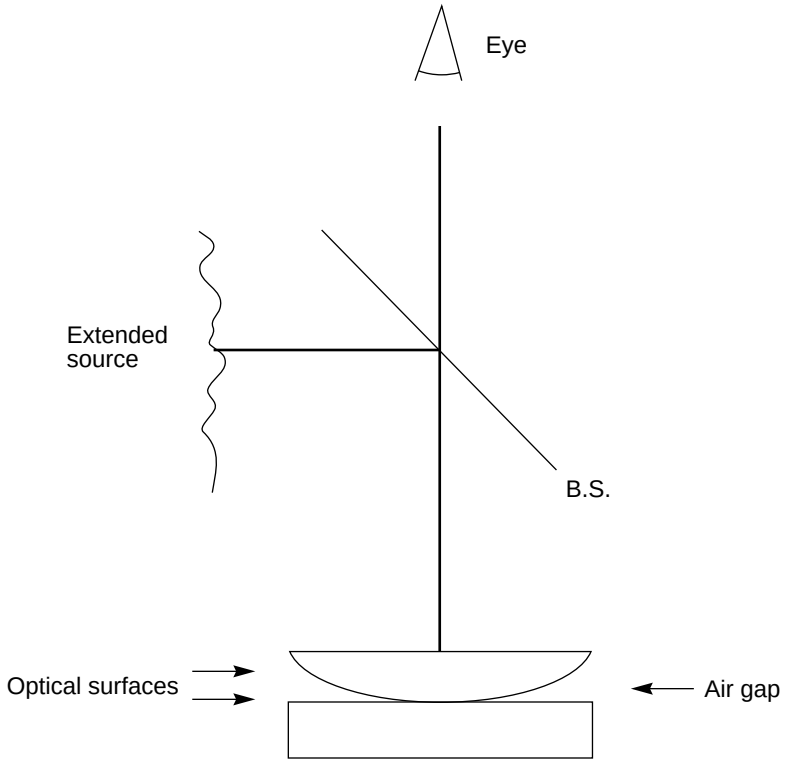


FIGURE 15 Fizeau interferometer.

are in contact at $\rho = 0$, the OPD between the reflected waves is twice the gap, and the condition for a dark fringe is

$$\rho = \sqrt{m\lambda R} \tag{54}$$

Circular fringes that increase in radius as the square root of ρ are observed. Note that a dark fringe occurs at the center of the pattern. In reflection, this point must be dark, as there is no interface at the contact point to produce a reflection.

Michelson Interferometer

There are many two-beam interferometers which allow the surfaces producing the two wavefronts to be physically separated by a large distance. These instruments allow the two wavefronts to travel along different optical paths. One of these is the *Michelson interferometer* diagramed in Fig. 16*a*. The two interfering wavefronts are produced by the reflections from the two mirrors. A plate beamsplitter with one face partially silvered is used, and an identical block of glass is placed in one of the arms of the interferometer to provide the same amount of glass path in each arm. This cancels the effects of the dispersion of the glass beamsplitter and allows the system to be used with white light since the optical path difference is the same for all wavelengths.

Figure 16*b* provides a folded view of this interferometer and shows the relative optical

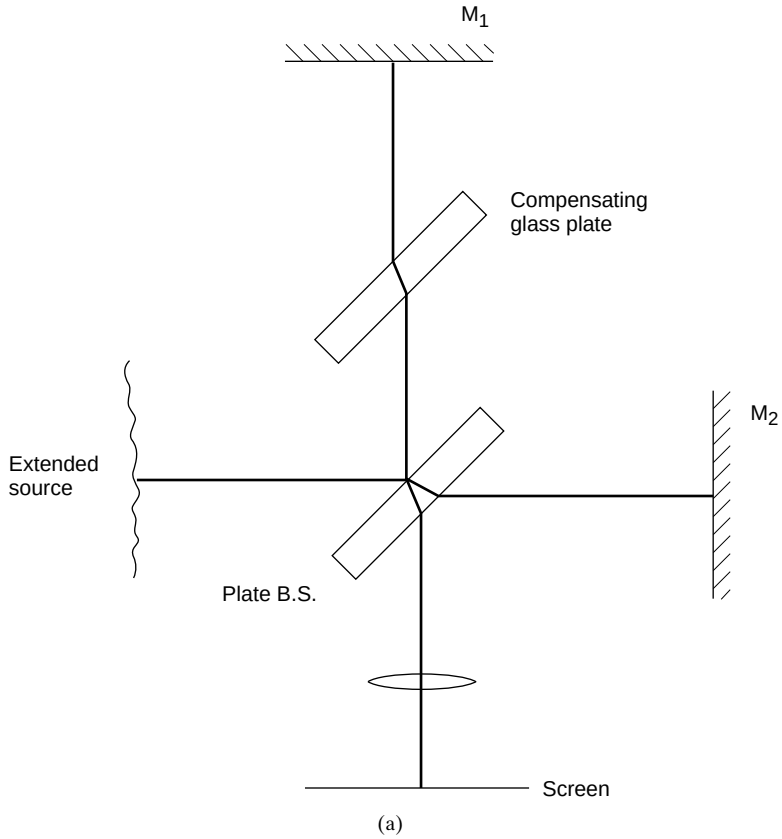
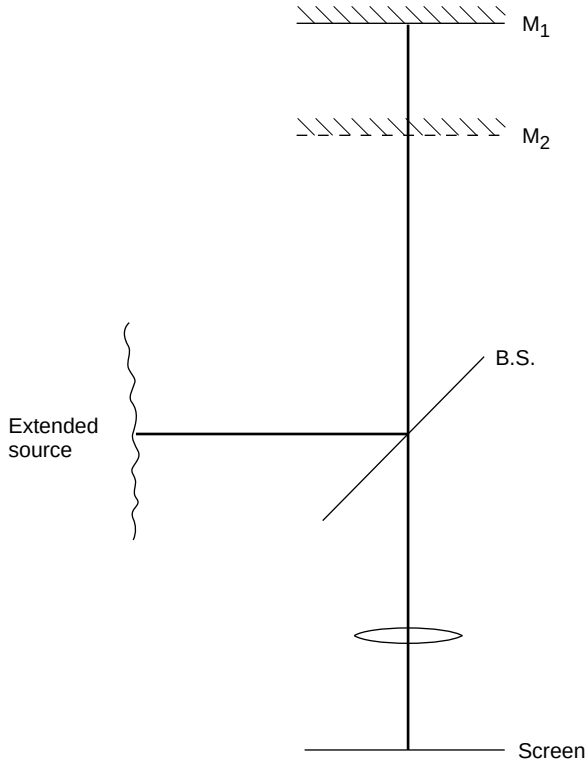


FIGURE 16 Michelson interferometer: (*a*) schematic view; and (*b*) folded view showing the relative optical position of the two mirrors.

position of the two mirrors as seen by the viewing screen. It should be obvious that the two mirrors can be thought of as the two surfaces of a "glass" plate that is illuminated by the source. In this case, the index of the fictitious plate is one, and the reflectivity at the two surfaces is that of the mirrors. Depending on the mirror orientations and shapes, the interferometer either mimics a plane-parallel plate of adjustable thickness, a wedge of arbitrary angle and thickness, or the comparison of a reference surface with an irregular or curved surface. The type of fringes that are produced will depend on this configuration, as well as on the source used for illumination.

When a monochromatic point source is used, nonlocalized fringes are produced, and the imaging lens is not needed. Two virtual-source images are produced, and the resulting fringes can be described by the interference of two spherical waves (discussed earlier). If the mirrors are parallel, circular fringes centered on the line normal to the mirrors result as with a plane-parallel plate. The source separation is given by twice the apparent mirror separation. If the mirrors have a relative tilt, the two source images appear to be laterally displaced, and hyperbolic fringes result. Along a plane bisecting the source images, straight equispaced fringes are observed.

When an extended monochromatic source is used, the interference fringes are localized. If the mirrors are parallel, fringes of equal inclination or Haidinger fringes (as described earlier) are produced. The fringes are localized at infinity and are observed in the rear



(b)

FIGURE 16 (*Continued*.)

focal plane of the imaging lens. Fringes of equal thickness localized at the mirrors are generated when the mirrors are tilted. The apparent mirror separation should be kept small, and the imaging lens should focus on the mirror surface.

If the extended source is polychromatic, colored fringes localized at the mirrors result. They are straight for tilted mirrors. The fringes will have high visibility only if the apparent mirror separation or OPD is smaller than the coherence length of the source. Another way of stating this is that the order of interference m must be small to view the colored fringes. As m increases, the fringes will wash out. The direct analogy here is a thin film. As the mirror separation is varied, the fringe visibility will vary. The fringe visibility as a function of mirror separation is related to the source frequency spectrum (see under ''Source Spectrum'' and ''Coherence and Interference''), and this interferometer forms the basis of a number of spectrometers. This topic is further discussed in Chap. 29, Vol. 2, ''Metrology.'' When the source spectrum is broad, chromatic fringes cannot be viewed with the mirrors parallel. This is because the order of interference for fringes of equal inclination is a maximum at the center of the pattern.

An important variation of the Michelson interferometer occurs when monochromatic collimated light is used. This is the *Twyman-Green interferometer,* and is a special case of point-source illumination with the source at infinity. Plane waves fall on both mirrors, and if the mirrors are flat, nonlocalized equispaced fringes are produced. Fringes of equal thickness can be viewed by imaging the mirrors onto the observation screen. If one of the mirrors is not flat, the fringes represent changes in the surface height. The two surfaces

are compared as in the Fizeau interferometer. This interferometer is an invaluable tool for optical testing.

2.7 MULTIPLE BEAM INTERFERENCE

Throughout the preceding discussions, we have assumed that only two waves were being interfered. There are many situations where multiple beams are involved. Two examples are the diffraction grating and a plane-parallel plate. We have been ignoring multiple reflections, and in some instances these extra beams are very important. The net electric field is the sum of all of the component fields. The two examples noted above present different physical situations: all of the interfering beams have a constant intensity with a diffraction grating, and the intensity of the beams from a plane-parallel plate decreases with multiple reflections.

Diffraction Grating

A *diffraction grating* can be modeled as a series of equispaced slits, and the analysis bears a strong similarity to the Young's double slit (discussed earlier). It operates by division of wavefront, and the geometry is shown in Fig. 17. The slit separation is d, the OPD between successive beams for a given observation angle θ is $d \sin(\theta)$, and the corresponding phase difference $\Delta\phi = 2\pi d \sin(\theta)/\lambda$. The field due to the nth slit at a distant observation point is

$$E_j(\theta) = Ae^{i(j-1)\Delta\phi} \qquad j = 1, 2, \ldots, N \tag{55}$$

where all of the beams have been referenced to the first slit, and there are N total slits. The net field is

$$E(\theta) = \sum_{j=1}^{N} E_j(\theta) = A\sum_{j=1}^{N} (e^{i\Delta\phi})^{j-1} \tag{56}$$

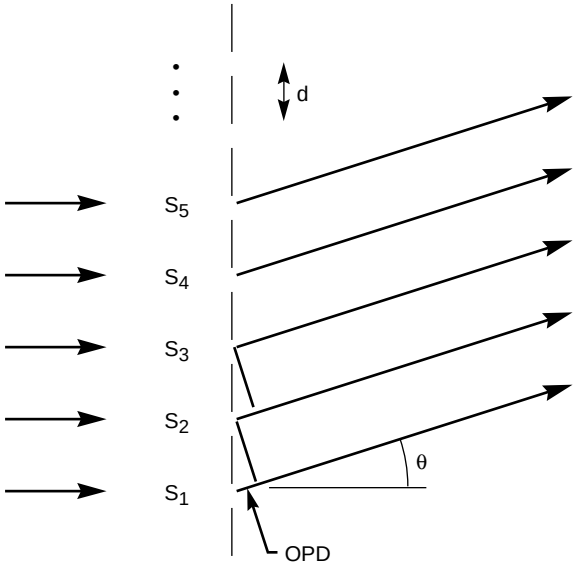


FIGURE 17 Diffraction grating: multiple-beam interference by division of wavefront.

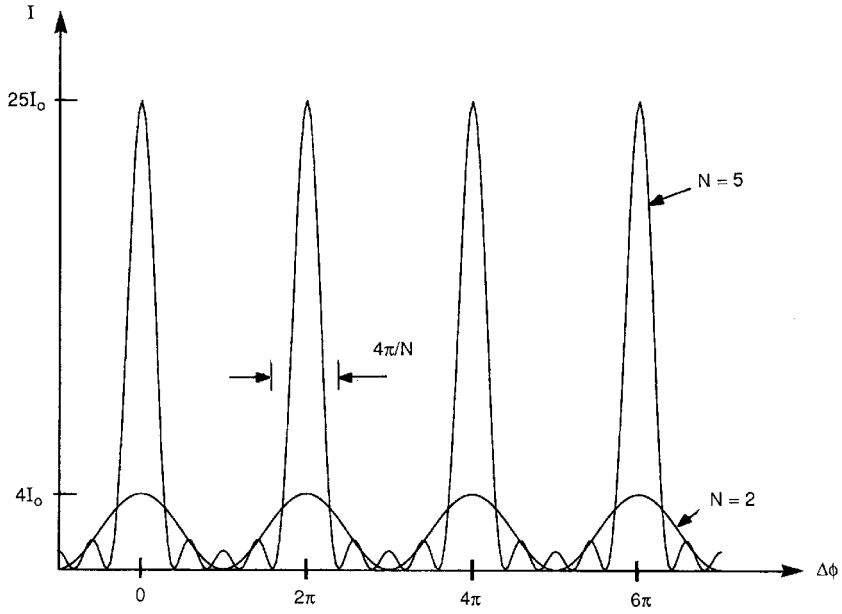


FIGURE 18 The interference patterns produced by gratings with 2 and 5 slits.

which simplifies to

$$E(\theta) = A\left(\frac{1 - e^{iN\Delta\phi}}{1 - e^{i\Delta\phi}}\right) \tag{57}$$

The resulting intensity is

$$I(\theta) = I_0\left[\frac{\sin^2\left(\dfrac{N\Delta\phi}{2}\right)}{\sin^2\left(\dfrac{\Delta\phi}{2}\right)}\right] = I_0\left[\frac{\sin^2\left(\dfrac{N\pi d\,\sin(\theta)}{\lambda}\right)}{\sin^2\left(\dfrac{\pi d\,\sin(\theta)}{\lambda}\right)}\right] \tag{58}$$

where I_0 is the intensity due to an individual slit.

This intensity pattern is plotted in Fig. 18 for $N = 5$. The result for $N = 2$, which is the double-slit experiment, is also shown. The first thing to notice is that the locations of the maxima are the same, independent of the number of slits. A maximum of intensity is obtained whenever the phase difference between adjacent slits is a multiple of 2π. These maxima occur at the diffraction angles given by

$$\sin(\theta) = \frac{m\lambda}{d} \tag{59}$$

where m is an integer. The primary difference between the two patterns is that with multiple slits, the intensity at the maximum increases to N^2 times that due to a single slit, and this energy is concentrated into a much narrower range of angles. The full width of a diffraction peak between intensity zero corresponds to a phase difference $\Delta\phi$ of $4\pi/N$.

The number of intensity zeros between peaks is $N-1$. As the number of slits increases, the angular resolution or resolving power of the grating greatly increases. The effects of a finite slit width can be added by replacing I_0 in Eq. (58) by the single-slit diffraction pattern. This intensity variation forms an envelope for the curve in Fig. 18.

Plane-Parallel Plate

The plane-parallel plate serves as a model to study the interference of multiple waves obtained by division of amplitude. As we shall see, the incremental phase difference between the interfering beams is constant but, in this case, the beams have different intensities. A plate of thickness t and index n with all of the reflected and transmitted beams is shown in Fig. 19. The amplitude reflection and transmission coefficients are ρ and ρ', and τ and τ', where the primes indicate reflection or transmission from within the plate. The first reflected beam is 180° out of phase with the other reflected beams since it is the only beam to undergo an external reflection, and $\rho = -\rho'$. Note that ρ' occurs only in odd powers for the reflected beams. Each successive reflected or transmitted beam is reduced in amplitude by ρ^2. The phase difference between successive reflected or transmitted beams is the same as we found when studying fringes of equal inclination from a plane-parallel plate:

$$\Delta\phi = \left[\frac{4\pi nt\cos(\theta')}{\lambda}\right] \tag{60}$$

where θ' is the angle inside the plate.

The transmitted intensity can be determined by first summing all of the transmitted amplitudes:

$$E(\Delta\phi) = \sum_{j=1}^{\infty} E_j = A\tau\tau' \sum_{j=1}^{\infty} (\rho^2 e^{i\Delta\phi})^{j-1} \tag{61}$$

where the phase is referenced to the first transmitted beam. The result of the summation is

$$E(\Delta\phi) = \left(\frac{A\tau\tau'}{1-\rho^2 e^{i\Delta\phi}}\right) \tag{62}$$

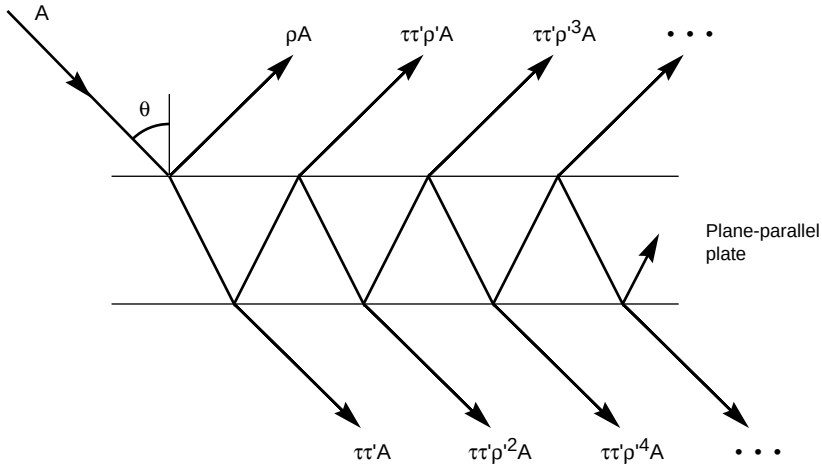


FIGURE 19 Plane-parallel plate: multiple-beam interference by division of amplitude.

The transmitted intensity I_t is the squared modulus of the amplitude which, after simplification, becomes

$$\frac{I_t}{I_0} = \frac{1}{1 + \left(\dfrac{2\rho}{1-\rho^2}\right)^2 \sin^2(\Delta\phi/2)} \tag{63}$$

where I_0 is the incident intensity. We have also assumed that there is no absorption in the plate, and therefore $\tau\tau' + \rho^2 = 1$. Under this condition of no absorption, the sum of the reflected and transmitted light must equal the incident light: $I_t + I_r = I_0$. The expressions for the transmitted and reflected intensities are then

$$\frac{I_t}{I_0} = \frac{1}{1 + F\sin^2(\Delta\phi/2)} \tag{64}$$

and

$$\frac{I_r}{I_0} = \frac{F\sin^2(\Delta\phi/2)}{1 + F\sin^2(\Delta\phi/2)} \tag{65}$$

and F is defined as

$$F \equiv \left(\frac{2\rho}{1-\rho^2}\right)^2 \tag{66}$$

F is the *coefficient of finesse* of the system and is a function of the surface reflectivity only. The value of F will have a large impact on the shape of the intensity pattern. Note that the reflected intensity could also have been computed by summing the reflected beams.

A maximum of transmitted intensity, or a minimum of reflected intensity, will occur when $\Delta\phi/2 = m\pi$, where m is an integer. Referring back to Eq. (60), we find that this corresponds to the angles

$$\cos\theta' = \frac{m\lambda}{2nt} \tag{67}$$

This is exactly the same condition that was found for a plane-parallel plate with two beams [Eq. (49)]. With an extended source, fringes of equal inclination are formed, and they are localized at infinity. They must be at infinity since all of the reflected or transmitted beams are parallel for a given input angle. The fringes are observed in the rear focal plane of a viewing lens. If the optical axis of this lens is normal to the surface, circular fringes about the axis are produced. The locations of the maxima and minima of the fringes are the same as were obtained with two-beam interference.

The shape of the intensity profile of these multiple beam fringes is not sinusoidal, as it was with two beams. A plot of the transmitted fringe intensity [Eq. (64)] as a function of $\Delta\phi$ is shown in Fig. 20 for several values of F. When the phase difference is a multiple of 2π, we obtain a bright fringe independent of F or ρ. When F is small, low-visibility fringes are produced. When F is large, however, the transmitted intensity is essentially zero unless the phase has the correct value. It drops off rapidly for even small changes in $\Delta\phi$. The transmitted fringes will be very narrow bright circles on an essentially black background. The reflected intensity pattern is one minus this result, and the fringe pattern will be very dark bands on a uniform bright background. The reflected intensity profile is plotted in Fig. 21 for several values of F.

The value of F is a strong function of the surface reflectivity $R = \rho^2$. We do not obtain appreciable values of F until the reflectivity is approximately one. For example, $R = 0.8$ produces $F = 80$, while $R = 0.04$ gives $F = 0.17$. This latter case is typical for uncoated glass, and dim broad fringes in reflection result, as in Fig. 21. The pattern is approximately sinusoidal, and it is clear that our earlier assumptions about ignoring multiple reflections when analyzing a plane-parallel plate are valid for many low-reflectivity situations.

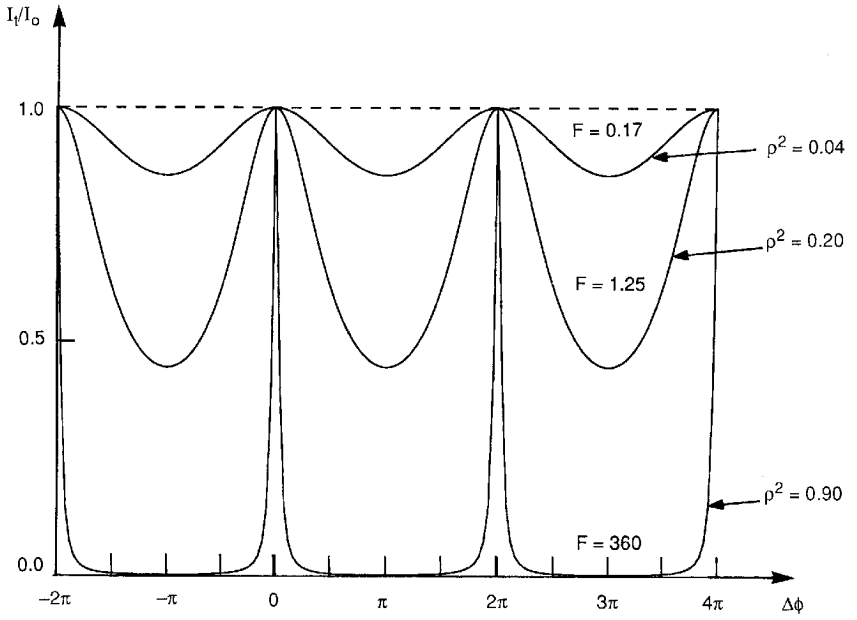


FIGURE 20 The transmitted intensity of a multiple-beam interference pattern produced by a plane-parallel plate.

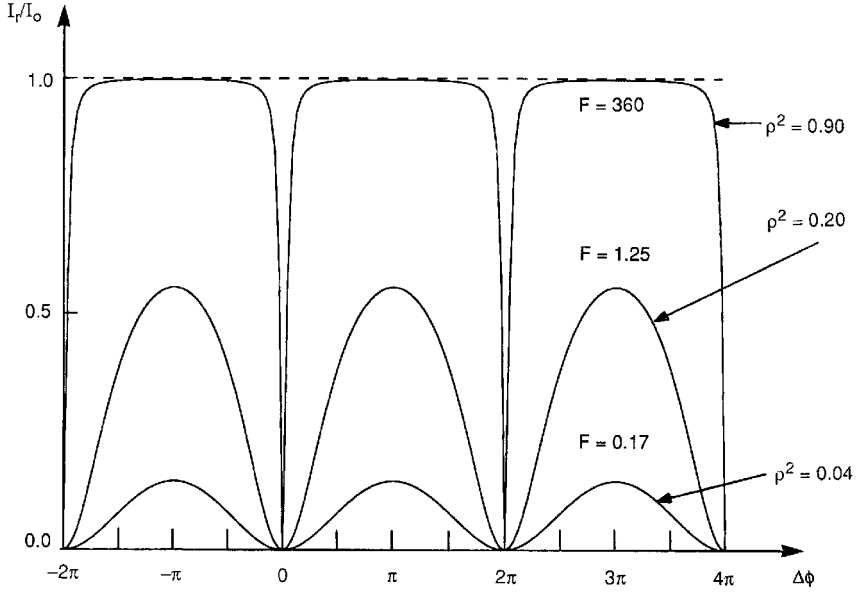


FIGURE 21 The reflected intensity of a multiple-beam interference pattern produced by a plane-parallel plate.

The multiple beam interference causes an energy redistribution much like that obtained from a diffraction grating. A strong response is obtained only when all of the reflected beams at a given angle add up in phase. The difference between this pattern and that of a diffraction pattern is that there are no oscillations or zeros between the transmitted intensity maxima. This is a result of the unequal amplitudes of the interfering beams. With a diffraction grating, all of the beams have equal amplitude, and the resultant intensity oscillates as more beams are added.

Multiple-beam fringes of equal thickness can be produced by two high-reflectivity surfaces in close proximity in a Fizeau interferometer configuration. The dark fringes will narrow to sharp lines, and each fringe will represent a contour of constant OPD between the surfaces. As before, a dark fringe corresponds to a gap of an integer number of half wavelengths. The area between the fringes will be bright. The best fringes will occur when the angle and the separation between the surfaces is kept small. This will prevent the multiple reflections from walking off or reflecting out of the gap.

Fabry-Perot Interferometer

The *Fabry-Perot interferometer* is an important example of a system which makes use of multiple-beam interference. This interferometer serves as a high-resolution spectrometer and also as an optical resonator. In this latter use, it is an essential component of a laser. The system is diagrammed in Fig. 22, and it consists of two highly reflective parallel surfaces separated by a distance t. These two separated reflective plates are referred to as a *Fabry-Perot etalon* or cavity, and an alternate arrangement has the reflected coatings applied to the two surfaces of a single glass plate. The two lenses serve to collimate the light from a point on the extended source in the region of the cavity and to then image this point onto the screen. The screen is located in the focal plane of the lens so that fringes of equal inclination localized at infinity are viewed. As we have seen, light of a fixed wavelength will traverse the etalon only at certain well-defined angles. Extremely sharp multiple-beam circular fringes in transmission are produced on the screen, and their profile is the same as that shown in Fig. 20.

If the source is not monochromatic, a separate independent circular pattern is formed for each wavelength. Equation (67) tells us that the location or scale of the fringes is dependent on the wavelength. If the source is composed of two closely spaced wavelengths, the ring structure is doubled, and the separation of the two sets of rings allows the hyperfine structure of the spectral lines to be evaluated directly. More

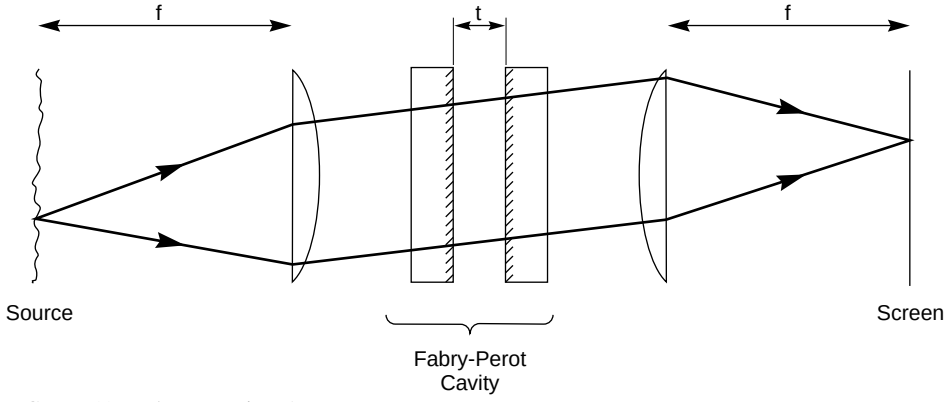


FIGURE 22 Fabry-Perot interferometer.

complicated spectra, usually composed of discrete spectral lines, can also be measured. This analysis is possible even though the order of interference is highest in the center of the pattern. If the phase change $\Delta\phi$ due to the discrete wavelengths is less than the phase change between adjacent fringes, nonoverlapping sharp fringes are seen.

A quantity that is often used to describe the performance of a Fabry-Perot cavity is the *finesse* $\mathcal{F}$. It is a measure of the number of resolvable spectral lines, and is defined as the ratio of the phase difference between adjacent fringes to the full width-half maximum FWHM of a single fringe. Since the fringe width is a function of the coefficient of finesse, the finesse itself is also a strong function of reflectivity. The phase difference between adjacent fringes is 2π, and the half width-half maximum can be found by setting Eq. (64) equal to $\frac{1}{2}$ and solving for $\Delta\phi$. The FWHM is twice this value, and under the assumption that F is large,

$$\text{FWHM} = \frac{4}{\sqrt{F}} \tag{68}$$

and the finesse is

$$\mathcal{F} \equiv \frac{2\pi}{\text{FWHM}} = \frac{\pi\sqrt{F}}{2} = \frac{\pi\rho}{1-\rho^2} = \frac{\pi\sqrt{R}}{1-R} \tag{69}$$

where ρ is the amplitude reflectivity, and R is the intensity reflectivity. Typical values for the finesse of a cavity with flat mirrors is about 30 and is limited by the flatness and parallelism of the mirrors. There are variations in $\Delta\phi$ across the cavity. Etalons consisting of two curved mirrors can be constructed with a much higher finesse, and values in excess of 10,000 are available.

Another way of using the Fabry-Perot interferometer as a spectrometer is suggested by rewriting the transmission [Eq. (64)] in terms of the frequency ν:

$$T = \frac{I_t}{I_0} = \frac{1}{1 + F\sin^2(2\pi t\nu/c)} \tag{70}$$

where Eq. (60) relates the phase difference to the wavelength, t is the mirror separation, and an index of one and normal incidence ($\theta' = 0$) have been assumed. This function is plotted in Fig. 23, and a series of transmission spikes separated in frequency by $c/2t$ are

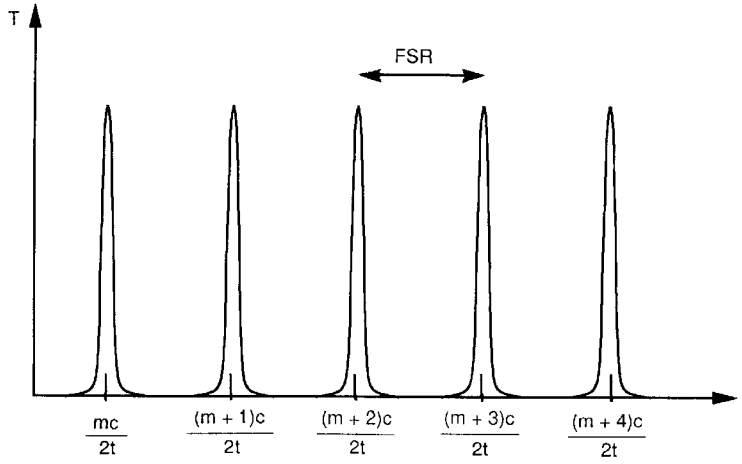


FIGURE 23 The transmission of a Fabry-Perot cavity as a function of frequency.

seen. A maximum occurs whenever the value of the sine is zero. The separation of these maxima is known as the *free spectral range,* FSR. If the separation of the mirrors is changed slightly, these transmission peaks will scan the frequency axis. Since the order of interference m is usually very large, it takes only a small change in the separation to move the peaks by one FSR. In fact, to scan one FSR, the required change in separation is approximately t/m. If the on-axis transmitted intensity is monitored while the mirror separation is varied, a high-resolution spectrum of the source is obtained. The source spectrum must be contained within one free spectral range so that the spectrum is probed by a single transmission peak at a time. If this were not the case, the temporal signal would contain simultaneous contributions from two or more frequencies resulting from different transmission peaks. Under this condition there are overlapping orders, and it is often prevented by using an auxiliary monochromator with the scanning Fabry-Perot cavity to preselect or limit the frequency range of the input spectrum. The resolution $\Delta\nu$ of the trace is limited by the finesse of the cavity.

For a specific cavity, the value of m at a particular transmission peak, and some physical insight into the operation of this spectrometer, is obtained by converting the frequency of a particular transmission mode $mc/2t$ into wavelength:

$$\lambda = \frac{2t}{m} \qquad \text{or} \qquad t = m\frac{\lambda}{2} \tag{71}$$

For the mth transmission maximum, exactly m half waves fit across the cavity. This also implies that the round-trip path within the cavity is an integer number of wavelengths. Under this condition, all of the multiply-reflected beams are in phase everywhere in the cavity, and therefore all constructively interfere. A maximum in the transmission occurs. Other maxima occur at different wavelengths, but these specific wavelengths must also satisfy the condition that the cavity spacing is an integer number of half wavelengths.

These results also allow us to determine the value of m. If a 1-cm cavity is used and the nominal wavelength is 500 nm, $m = 40{,}000$ and $\text{FSR} = 1.5 \times 10^{10}$ Hz. The wavelength interval corresponding to this FSR is 0.0125 nm. If a 1-mm cavity is used instead, the results are $m = 4000$ and $\text{FSR} = 1.5 \times 10^{11}$ Hz $= 0.125$ nm. We see now that to avoid overlapping orders, the spectrum must be limited to a very narrow range, and this range is a function of the spacing. Cavities with spacings of a few tens of μm's are available to increase the FSR. Increasing the FSR does have a penalty. The finesse of a cavity depends only on the reflectivities, so as the FSR is increased by decreasing t, the FWHM of the transmission modes increases to maintain a constant ratio. The number of resolvable spectrum lines remains constant, and the absolute spectral resolution decreases.

A mirror translation of a half wavelength is sufficient to cover the FSR of the cavity. The usual scanning method is to separate the two mirrors with a piezoelectric spacer. As the applied voltage is changed, the cavity length will also change. An alternate method is to change the index of the air in the cavity by changing the pressure.

2.8 COHERENCE AND INTERFERENCE

The observed fringe visibility is a function of the spatial and temporal coherence of the source. The classical assumption for the analysis is that every point on an extended source radiates independently and therefore produces its own interference pattern. The net intensity is the sum of all of the individual intensity patterns. In a similar manner, each wavelength or frequency of a nonmonochromatic source radiates independently, and the

temporal average is the sum of the individual temporal averages. *Coherence theory* allows the interference between the light from two point sources to be analyzed, and a good visual model is an extended source illuminating the two pinholes in Young's double slit. We need to determine the relationship between the light transmitted through the two pinholes. Coherence theory also accounts for the effects of the spectral bandwidth of the source.

With interference by division of amplitude using an extended source, the light from many point sources is combined at the observation point, and the geometry of the interferometer determines where the fringes are localized. Coherence theory will, however, predict the spectral bandwidth effects for division of amplitude interference. Each point on the source is interfered with an image of that same point. The temporal coherence function relates the interference of these two points independently of other points on the source. The visibility function for the individual interference pattern due to these two points is computed, and the net pattern is the sum of these patterns for the entire source. The temporal coherence effects in division of amplitude interference are handled on a point-by-point basis across the source.

In this section, the fundamentals of coherence theory as it relates to interference are introduced. Much more detail on this subject can be found in Chap. 4, "Coherence Theory."

Mutual Coherence Function

We will consider the interference of light from two point sources or pinholes. This light is derived from a common origin so that there may be some relationship between the complex fields at the two sources. We will represent these amplitudes at the pinholes as $E_1(t)$ and $E_2(t)$, as shown in Fig. 24. The propagation times between the two sources and the observation point are t_1 and t_2, where the times are related to the optical path lengths by $t_i = \mathrm{OPL}_i/c$. The two complex amplitudes at the observation point are then $E_1(t - t_1)$ and $E_2(t - t_2)$, where the amplitudes have been scaled to the observation plane. The time-average intensity at the observation point can be found by returning to Eq. (13), which is repeated here with the time dependence:

$$I = I_1 + I_2 + \langle E_1(t - t_1)E_2^*(t - t_2)\rangle + \langle E_1^*(t - t_1)E_2(t - t_2)\rangle \tag{72}$$

where I_1 and I_2 are the intensities due to the individual sources. If we now shift our time origin by t_2, we obtain

$$I = I_1 + I_2 + \langle E_1(t + \tau)E_2^*(t)\rangle + \langle E_1^*(t + \tau)E_2(t)\rangle \tag{73}$$

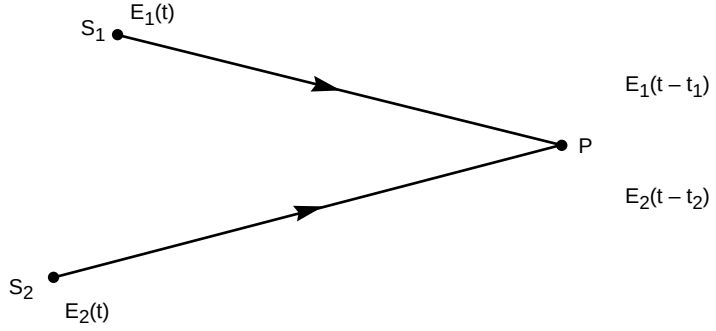


FIGURE 24 Geometry for examining the mutual coherence of two sources.

where

$$\tau = t_2 - t_1 = \frac{\text{OPL}_2 - \text{OPL}_1}{c} = \frac{\text{OPD}}{c} \tag{74}$$

The difference in transit times for the two paths is τ. The last two terms in the expression for the intensity are complex conjugates, and they contain the interference terms.

We will now define the *mutual coherence function* $\Gamma_{12}(\tau)$:

$$\Gamma_{12}(\tau) = \langle E_1(t+\tau)E_2^*(t)\rangle \tag{75}$$

which is the cross correlation of the two complex amplitudes. With this identification, the intensity of the interference pattern is

$$I = I_1 + I_2 + \Gamma_{12}(\tau) + \Gamma_{12}^*(\tau) \tag{76}$$

or, recognizing that a quantity plus its complex conjugate is twice the real part,

$$I = I_1 + I_2 + 2\,\text{Re}\,\{\Gamma_{12}(\tau)\} \tag{77}$$

It is convenient to normalize the mutual coherence function by dividing by the square root of the product of the two self-coherence functions. The result is the *complex degree of coherence*:

$$\gamma_{12}(\tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{\Gamma_{11}(0)\Gamma_{22}(0)}} = \frac{\Gamma_{12}(\tau)}{\sqrt{\langle|E_1(t)|^2\rangle\langle|E_2(t)|^2\rangle}} = \frac{\Gamma_{12}(\tau)}{\sqrt{I_2 I_2}} \tag{78}$$

and the intensity can be rewritten:

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\,\text{Re}\,\{\gamma_{12}(\tau)\} \tag{79}$$

We can further simplify the result by writing $\gamma_{12}(\tau)$ as a magnitude and a phase:

$$\gamma_{12}(\tau) = |\gamma_{12}(\tau)|\, e^{i\phi_{12}(\tau)} = |\gamma_{12}(\tau)|\, e^{i[\alpha_{12}(\tau) - \Delta\phi(\tau)]} \tag{80}$$

where $\alpha_{12}(\tau)$ is associated with the source, and $\Delta\phi(\tau)$ is the phase difference due to the OPD between the two sources and the observation point [Eq. (18)]. The quantity $|\gamma_{12}(\tau)|$ is known as the *degree of coherence.* The observed intensity is therefore

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\,|\gamma_{12}(\tau)|\cos\,[\alpha_{12}(\tau) - \Delta\phi(\tau)] \tag{81}$$

The effect of $\alpha_{12}(\tau)$ is to add a phase shift to the intensity pattern. The fringes will be shifted. A simple example of this situation is Young's double-slit experiment illuminated by a tilted plane wave or a decentered source. With quasi-monochromatic light, the variations of both $|\gamma_{12}(\tau)|$ and $\alpha_{12}(\tau)$ with τ are slow with respect to changes of $\Delta\phi(\tau)$, so that the variations in the interference pattern in the observation plane are due primarily to changes in $\Delta\phi$ with position.

A final rewrite of Eq. (81) leads us to the intensity pattern at the observation point:

$$I = I_0\left\{1 + \frac{2\sqrt{I_1 I_2}}{I_1 + I_2}|\gamma_{12}(\tau)|\cos\,[\alpha_{12}(\tau) - \Delta\phi(\tau)]\right\} \tag{82}$$

where $I_0 = I_1 + I_2$. The fringe visibility is therefore

$$\gamma(\tau) = \frac{2\sqrt{I_1 I_2}}{I_1 + I_2}|\gamma_{12}(\tau)| \tag{83}$$

and is a function of the degree of coherence and τ. Remember that τ is just the temporal

measure of the OPD between the two sources and the observation point. If the two intensities are equal, the fringe visibility is simply the degree of coherence: $\gamma(\tau) = |\gamma_{12}(\tau)|$. The degree of coherence will take on values betwen 0 and 1. The source is *coherent* when $|\gamma_{12}(\tau)| = 1$, and completely incoherent when $|\gamma_{12}(\tau)| = 0$. The source is said to be *partially coherent* for other values. No fringes are observed with an incoherent source, and the visibility is reduced with a partially coherent source.

Spatial Coherence

The spatial extent of the source and its distance from the pinholes will determine the visibility of the fringes produced by the two pinhole sources (see Fig. 25). Each point on the source will produce a set of Young's fringes, and the position of this pattern in the observation plane will shift with source position. The value of $\alpha_{12}(\tau)$ changes with source position. The existence of multiple shifted patterns will reduce the overall visibility. As an example, consider a quasi-monochromatic source that consists of a several point sources arranged in a line. Each produces a high modulation fringe pattern in the observation plane (Fig. 26*a*), but there is a lateral shift between each pattern. The net pattern shows a fringe with the same period as the individual patterns, but it has a reduced modulation due to the shifts (Fig. 26*b*). This reduction in visibility can be predicted by calculating the degree of coherence $|\gamma_{12}(\tau)|$ at the two pinholes.

Over the range of time delays between the interfering beams that are usually of interest, the degree of coherence is a slowly varying function and is approximately equal to the value at $\tau = 0$: $|\gamma_{12}(\tau)| = |\gamma_{12}(0)| = |\gamma_{12}|$. The *van Cittert–Zernike theorem* allows the degree of coherence in the geometry of Fig. 25 to be calculated. Let θ be the angular separation of the two pinholes as seen from the source. This theorem states that degree of coherence between two points is the modulus of the scaled and normalized Fourier transform of the source intensity distribution:

$$|\gamma_{12}| = \left| \frac{\iint_S I(\xi, \eta) e^{i(2\pi/\lambda)(\xi\theta_x + \eta\theta_y)} d\xi\, d\eta}{\iint_S I(\xi, \eta)\, d\xi\, d\eta} \right| \tag{84}$$

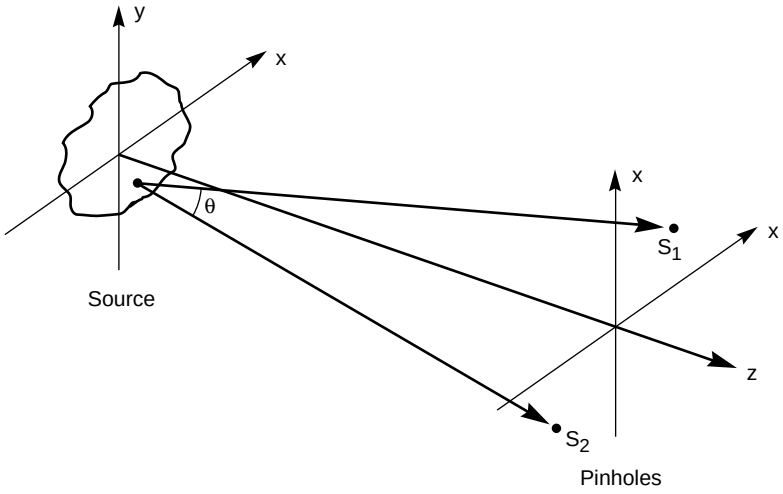


FIGURE 25 An extended source illuminating two pinholes.

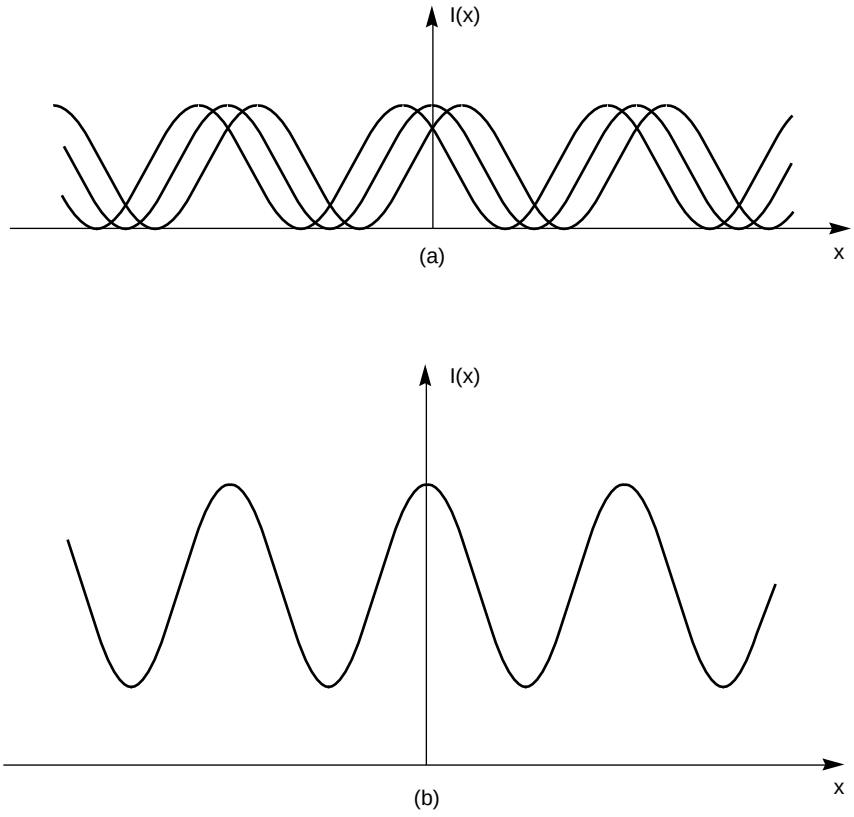


FIGURE 26 The interference pattern produced by a linear source: (*a*) the individual fringe patterns; and (*b*) the net fringe pattern with reduced visibility.

where θ_x and θ_y are the x and y components of the pinhole separation θ, and the integral is over the source.

Two cases that are of particular interest are a slit source and a circular source. The application of the van Cittert–Zernike theorem yields the two coherence functions:

$$\text{Slit source of width } w: \quad |\gamma_{12}| = \left|\text{sinc}\left(\frac{w\theta_x}{\lambda}\right)\right| = \left|\text{sinc}\left(\frac{wa}{\lambda z}\right)\right| \tag{85}$$

$$\text{Circular source of diameter } d: \quad |\gamma_{12}| = \left|\frac{2J_1\left(\dfrac{\pi d\theta_x}{\lambda}\right)}{\dfrac{\pi d\theta_x}{\lambda}}\right| = \left|\frac{2J_1\left(\dfrac{\pi da}{\lambda z}\right)}{\dfrac{\pi da}{\lambda z}}\right| \tag{86}$$

where a is the separation of the pinholes, z is the distance from the source to the pinholes, the sinc function is defined by Eq. (42), and J_1 is a first-order Bessel function. The pinholes are assumed to be located on the x-axis. These two functions share the common characteristic of a central core surrounded by low-amplitude side lobes. We can imagine these functions of pinhole spacing mapped onto the aperture plane. The coherence

function is centered on one of the pinholes. If the other pinhole is then within the central core, high-visibility fringes are produced. If the pinhole spacing places the second pinhole outside the central core, low-visibility fringes result.

Michelson Stellar Interferometer

The *Michelson stellar interferometer* measures the diameter of stars by plotting out the degree of coherence due to the light from the star. The system is shown in Fig. 27. Two small mirrors separated by the distance a sample the light and serve as the pinholes. The spacing between these mirrors can be varied. This light is then directed along equal path lengths into a telescope, and the two beams interfere in the image plane. To minimize chromatic effects, the input light should be filtered to a small range of wavelengths. The modulation of the fringes is measured as a function of the mirror spacing to measure the degree of coherence in the plane of the mirrors. This result will follow Eq. (86) for a circular star, and the fringe visibility will go to zero when $a = 1.22\lambda\alpha$, where $\alpha = d/z$ is the angular diameter of the star. We measure the mirror separation that produces zero visibility to determine α. In a similar manner, this interferometer can be used to measure the spacing of two closely spaced stars.

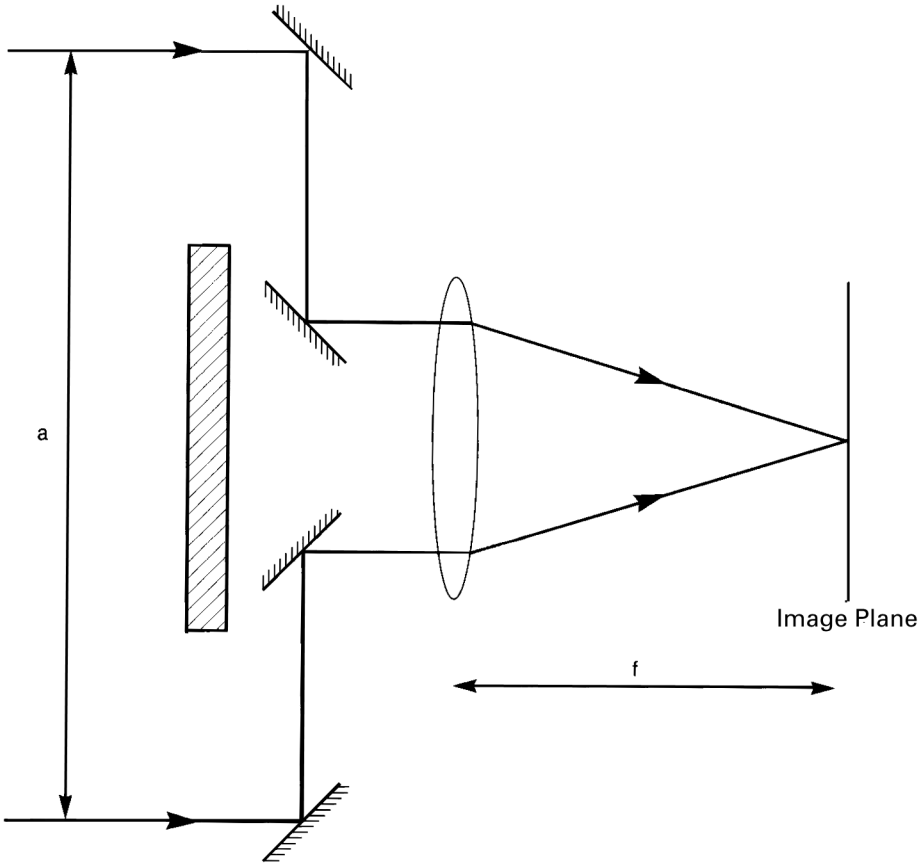


FIGURE 27 Michelson stellar interferometer.

Temporal Coherence

When examining temporal coherence effects, we use a source of small dimensions (a point source) that radiates over a range of wavelengths. The light from this source is split into two beams and allowed to interfere. One method to do this is to use an amplitude-splitting interferometer. Since the two sources are identical, the mutual coherence function becomes the *self-coherence function* $\Gamma_{11}(\tau)$. Equal-intensity beams are assumed. The complex degree of temporal coherence becomes

$$\gamma_{11}(\tau) = \frac{\Gamma_{11}(\tau)}{\Gamma_{11}(0)} = \frac{\langle E_1(t+\tau)E_1^*(t)\rangle}{\langle |E_1(t)|^2\rangle} \tag{87}$$

After manipulation, it follows from this result that $\gamma_{11}(\tau)$ is the normalized Fourier transform of the source intensity spectrum $S(\nu)$:

$$\gamma_{11}(\tau) = \frac{FT\{S(\nu)\}}{\int_0^\infty S(\nu)\,d\nu} = \frac{\int_0^\infty S(\nu)e^{i2\pi\nu\tau}\,d\nu}{\int_0^\infty S(\nu)\,d\nu} \tag{88}$$

The fringe visibility is the modulus of this result. Since $\gamma_{11}(\tau)$ has a maximum at $\tau = 0$, the maximum fringe visibility will occur when the time delay between the two beams is zero. This is consistent with our earlier observation under ''Source Spectrum'' that the fringes will be localized in the vicinity of zero OPD.

As an example, we will repeat the earlier problem of a uniform source spectrum:

$$S(\nu) = \text{rect}\left(\frac{\nu - \nu_0}{\Delta\nu}\right) \tag{89}$$

where ν_0 is the average frequency and $\Delta\nu$ is the bandwidth. The resulting intensity pattern is

$$I = I_0\{1 + \text{Re}\,\{\gamma_{12}(\tau)\}\} = I_0[1 + \text{sinc}\,(\tau\Delta\nu)\cos(2\pi\tau\nu_0)] \tag{90}$$

where the sinc function is the Fourier transform of the rect function. Using $\tau = \text{OPD}/c$ from Eq. (74), we can rewrite this equation in terms of the OPD to obtain the same result expressed in Eq. (47).

Laser Sources

The laser is an important source for interferometry, as it is a bright source of coherent radiation. Lasers are not necessarily monochromatic, as they may have more than one longitudinal mode, and it is important to understand the unique temporal coherence properties of a laser in order to get good fringes. The laser is a Fabry-Perot cavity that contains a gain medium. Its output spectrum is therefore a series of discrete frequencies separated by $c/2nL$, where L is the cavity length. For gas lasers, the index is approximately equal to one, and we will use this value for the analysis. If $G(\nu)$ is the gain bandwidth, the frequency spectrum is

$$S(\nu) = G(\nu)\,\text{comb}\left(\frac{2L\nu}{c}\right) \tag{91}$$

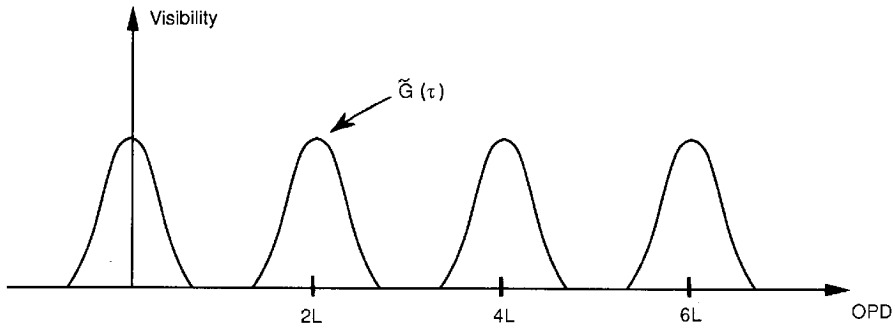


FIGURE 28 The fringe visibility versus OPD for a laser source.

where a comb function is a series of equally spaced delta functions. The number of modes contained under the gain bandwidth can vary from 1 or 2 up to several dozen. The resulting visibility function can be found by using Eq. (88):

$$\gamma(\tau) = |\gamma_{11}(\tau)| = \left|\tilde{G}(\tau) * \text{comb}\left(\frac{c\tau}{2L}\right)\right| = \left|\tilde{G}(\tau) * \text{comb}\left(\frac{\text{OPD}}{2L}\right)\right| \tag{92}$$

where $\tilde{G}(\tau)$ is the normalized Fourier transform of the gain bandwidth, and $*$ indicates convolution. This result is plotted in Fig. 28, where $\tilde{G}(\tau)$ is replicated at multples of $2L$. The width of these replicas is inversely proportional to the gain bandwidth. We see that as long as the OPD between the two optical paths is a multiple of twice the cavity length, high-visibility fringes will result. This condition is independent of the number of longitudinal modes of the laser. If the laser emits a single frequency, it is a coherent source and good visibility results for any OPD.

2.9 REFERENCES

M. Born and E. Wolf, *Principles of Optics,* Pergamon, New York, 1975.

R. W. Ditchburn, *Light,* Academic, New York, 1976.

M. Francon, *Optical Interferometry,* Academic, New York, 1966.

M. H. Freeman, *Optics,* Butterworths, London, 1990.

J. D. Gaskill, *Linear Systems, Fourier Transforms and Optics,* Wiley, New York, 1978.

P. Hariharan, *Optical Interferometry,* Academic Press, San Diego, Calif., 1985.

P. Hariharan (ed.), *Selected Papers on Interferometry,* SPIE, Bellingham, Wash., 1990.

P. Hariharan, *Basics of Interferometry,* Academic Press, San Diego, Calif., 1992.

E. Hecht, *Optics,* Addison-Wesley, Reading, Mass., 1989.

F. A. Jenkins and H. E. White, *Fundamentals of Optics,* McGraw-Hill, New York, 1976.

R. A. Longhurst, *Geometrical and Physical Optics,* Longman, London, 1973.

D. Malacara (ed.), *Selected Papers on Optical Shop Metrology,* SPIE, Bellingham, Wash., 1990.

D. Malacara (ed.), *Optical Shop Testing,* 2d edition, Wiley, New York, 1992.

A. S. Marathay, *Elements of Optical Coherence Theory,* Wiley, New York, 1982.

J. R. Meyer-Arendt, *Introduction to Classical and Modern Optics,* Prentice-Hall, Englewood Cliffs, N.J., 1989.

G. O. Reynolds, J. B. De Velis, G. B. Parrent, and B. J. Thompson, *The New Physical Optics Notebook: Tutorials in Fourier Optics,* SPIE, Bellingham, Wash., 1989.
B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics,* Wiley, New York, 1991.
W. H. Steel, *Interferometry,* Cambridge, London, 1967.
F. G. Smith and J. H. Thompson, *Optics,* Wiley, New York, 1971.
R. W. Wood, *Physical Optics,* Optical Society of America, Washington D.C., 1988.